P.B.5818 – Patentlaan 2
2280 HV Rijswijk (ZH)
☎ +31 70 340 2040
TX 31651 epo nl
FAX +31 70 340 3016

Europäisches Patentamt
Zweigstelle in Den Haag Recherchen-abteilung

European Patent Office
Branch at The Hague Search division

Office européen des brev ts
Département à La Haye Division de la recherche

De Clercq, Ann
Ann De Clercq & Co B.V.B.A.,
E. Gevaertdreef 10a
9830 Sint-Martens-Latem
BELGIQUE

Datum/Date
14.08.00

Zeichen/Ref./Réf.	Anmeldung Nr./Application No./Demande n°./Patent Nr./Patent No./Brevet n°.
INNX/35-EPDIV1	99118785.7-2105-

Anmelder/Applicant/Demandeur/Patentinhaber/Proprietor/Titulaire
INNOGENETICS N.V.

COMMUNICATION

The European Patent Office herewith transmits as an enclosure the European search report for the above-mentioned European patent application.

If applicable, copies of the documents cited in the European search report are attached.

☐ Additional set(s) of copies of the documents cited in the European search report is (are) enclosed as well.

The following specifications given by the applicant have been approved by the Search Division:

☒ abstract ☐ title

☐ The abstract was modified by the Search Division and the definitive text is attached to this communication.

The following figure will be published together with the abstract: NONE

REFUND OF THE SEARCH FEE

If applicable under Article 10 Rules relating to fees, a separate communication from the Receiving Section on the refund of the search fee will be sent later.

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 99 11 8785

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	S. MORI ET AL.: "A new type of hepatitis C in patients in Thailand" BIOCHEM. BIOPHYS. RES. COMMUN., vol. 183, no. 1, 1992, pages 334-342, XP000371000 * figure 1 *	1-21	C12N15/40 C12Q1/68 A61K39/29 G01N33/576 C12Q1/70 C07K14/18 C07K16/10
X,P	L. STUYVER ET AL.: "Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3" BIOCHEM. BIOPHYS. RES. COMMUN., vol. 192, no. 2, 1993, pages 635-641, XP000378704 * figure 1 *	1-21	A61P31/14
X	N. KATO ET AL.: "Molecular cloning of the human hepatitis C virus genome form Japanese patients with non-A non-B hepatitis" PROC. NATL. ACAD. SCI. USA, vol. 87, 1990, pages 9254-9258, XP000168621 * figure 2 *	1-21	
X,P	WO 93 10239 A (COMMON SERVICES AGENCY) 27 May 1993 (1993-05-27) * figure 9A *	1-21	
X	WO 92 19743 A (CHIRON CORPORATION) 12 November 1992 (1992-11-12) * figure 2 *	1-21	

-/--

TECHNICAL FIELDS SEARCHED (Int.Cl.7)

C07K
C12N
C12Q
A61K
G01N

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	2 August 2000	Skelly, J

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 99 11 8785

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X,P	DATABASE GENBAN 'Online! Accession No. X78863, 20 May 1994 (1994-05-20) VAN DOORN ET AL.: "Sequence analysis of hepatitis C virus genotypes 1 to 5" XP002017147 * abstract * & J. GEN. VIROL., vol. 76, 1994, pages 1871-1876,	1-21	
X,P	DATABASE GENBAN 'Online! Accession No D26387, 4 February 1994 (1994-02-04) HOTTA ET AL.: "Subtype analysis of hepatitis C virus in Indonesia" XP002017146 * abstract * & J. CLIN. MICROBIOL., vol. 32, 1994, pages 3049-3051,	1-21	

TECHNICAL FIELDS SEARCHED (Int.Cl.7)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	2 August 2000	Skelly, J

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 99 11 8785

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in th European Patent Office EDP file on
The European Patent Office is in no way liable for th se particulars which are merely given for the purpose of information.

02-08-2000

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9310239 A	27-05-1993	AU 671967 B	19-09-1996
		AU 3088792 A	15-06-1993
		CA 2123875 A	27-05-1993
		DE 610436 T	03-08-1995
		EP 0610436 A	17-08-1994
		ES 2065863 T	01-03-1995
		FI 942369 A	19-07-1994
		JP 7501442 T	16-02-1995
		US 5763159 A	09-06-1998
		ZA 9209015 A	17-05-1993
WO 9219743 A	12-11-1992	AU 668355 B	02-05-1996
		AU 2155892 A	21-12-1992
		BG 62142 B	31-03-1999
		BG 98200 A	31-01-1995
		BG 101876 A	30-11-1998
		CA 2108466 A	09-11-1992
		CZ 9601210 A	14-08-1996
		CZ 9302377 A	13-04-1994
		EP 0585398 A	09-03-1994
		FI 934937 A	05-01-1994
		HU 69609 A	28-09-1995
		JP 6508026 T	14-09-1994
		NO 934019 A	05-11-1993
		PL 169880 B	30-09-1996
		PL 170151 B	31-10-1996
		PT 100472 A,B	31-08-1993
		SK 123293 A	08-06-1994
		US 6071693 A	06-06-2000

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	FOR FURTHER ACTION	see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.
PCT.9546.HCV		
International application No.	International filing date (day/month/year)	(Earliest) Priority Date (day/month/year)
PCT/EP 95/04155	23/10/95	21/10/94
Applicant		
INNOGENETICS N.V. et al.		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 5 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. ☒ **Certain claims were found unsearchable** (see Box I).

2. ☒ **Unity of invention is lacking** (see Box II).

3. ☒ The international application contains disclosure of a **nucleotide and/or amino acid sequence listing** and the international search was carried out on the basis of the sequence listing
   - ☐ filed with the international application.
   - ☒ furnished by the applicant separately from the international application,
     - ☐ but not accompanied by a statement to the effect that it did not include matter going beyond the disclosure in the international application as filed.
   - ☐ Transcribed by this Authority

4. With regard to the **title,**
   - ☒ the text is approved as submitted by the applicant.
   - ☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**
   - ☒ the text is approved as submitted by the applicant.
   - ☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is:
   Figure No. _____
   - ☐ as suggested by the applicant.
   - ☐ because the applicant failed to suggest a figure.
   - ☐ because this figure better characterizes the invention.

   ☒ None of the figures.

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 37,38,50,51,61,62
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 37,38,50,51,61 and 62 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the claimed compositions.

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. Claims: 1-62 (partly and as far as applicable)
2.-14. Claims: 1-62 (partly and as far as applicable)

see continuation-sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-62 (partly and as far as applicable)

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/210

1. Claims 1-62 (partly and as far as applicable):
Nucleic acids from an HCV subtype 1d sequence, use thereof for diagnostic purposes, peptides encoded and corresponding antibodies and kits containing primers or probes derived from HCV subtype 1d.

2-14. Claims 1-62 (partly and as far as applicable):
As above, from HCV subtypes 1e, 1f, 2e, 2f, 2g, 2h, 2i, 4k, 41 or HCV types 7, 8, 9 or 10.

The subject-matter of invention 1 has been searched, whereas no search has been performed with regard to the subject-matter of inventions 2 to 14.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 95/04155

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/40 C07K14/18 C12Q1/70 C07K16/10 G01N33/569
A61K39/29

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K C12Q G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GENE, vol.114, no.2, 15 May 1992, AMSTERDAM,NL pages 245 - 250 K.LIU ET AL. 'Genomic typing of hepatitis C viruses present in China' see figure 2 ---	1-26, 30-35, 41-45
X	JOURNAL OF GENERAL VIROLOGY, vol.74, no.6, 3, READING GB pages 1093 - 1102 L.STUYVER ET AL. 'Typing of hepatitis C virus isolates and characterization of new subtypes using a line probe assay' see tables 1,2 --- -/--	1-26, 30-35, 41-45

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) .

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

3

Date of the actual completion of the international search	Date of mailing of the international search report
11 April 1996	02.07.96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	CUPIDO, M

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol.91, no.21, 11 June 1994, WASHINGTON pages 10134 - 10138 STUYVER ET AL. 'Classification of hepatitis C viruses based on phylogenetic analysis of the envelope 1 and nonstructural 5B regions and identification of five additional subtypes' see figures 1,2 ---	1-26, 30-35, 41-45
X	JOURNAL OF HEPATOLOGY, vol.21, no.1, COPENHAGEN, DK pages 122 - 129 VAN DOORN ET AL. 'Analysis of hepatitis C virus genotypes by a line probe assay and correlation with antibody profiles' see the whole document ---	1-26, 30-35, 41-45
X	WO,A,93 00365 (CHIRON CORPORATION) 7 January 1993 see page 34, line 19 ---	22-62
X	WO,A,93 06126 (CHIRON CORPORATION) 1 April 1993 see figures 2-1 ---	22-62
P,X	WO,A,94 25601 (N.V. INNOGENETICS S.A.) 10 November 1994 see figure 1 sequences EG-13 and EG-19 from HCV subtype 4k -----	1-62

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 95/04155

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9300365	07-01-93	AU-B-	2305392	25-01-93
		BG-A-	98332	28-02-95
		CA-A-	2110058	07-01-93
		EP-A-	0591431	13-04-94
		FI-A-	935808	21-02-94
		JP-T-	6508837	06-10-94
		NO-A-	934542	10-02-94
WO-A-9306126	01-04-93	AU-B-	2643692	27-04-93
		BG-A-	98653	31-05-95
		CA-A-	2116764	01-04-93
		EP-A-	0608261	03-08-94
		FI-A-	941199	27-04-94
		HU-A-	67342	28-03-95
		JP-T-	6511149	15-12-94
WO-A-9425601	10-11-94	AU-B-	6722294	21-11-94
		CA-A-	2139100	10-11-94
		CN-A-	1108030	06-09-95
		EP-A-	0651807	10-05-95
		FI-A-	946066	23-12-94
		JP-T-	7508423	21-09-95
		NO-A-	944967	21-12-94

P.B.5818 - Patentlaan
2280 HV Rijswijk (ZH)
(070) 3 40 20 40
31651 epo nl
(070) 3 40 30 16

Europäisches Patentamt
Zweigstelle in Den Haag Recherchenabteilung

European Patent Office
Branch at The Hague Search division

Office européen des brevets
Département à La Haye Division de la recherche

Grosset-Fournier, Chantal Catherine
Grosset-Fournier & Demachy s.a.r.l.
103 rue La Fayette
75010 Paris
FRANCE

Datum/Date
17. 11. 96

Zeichen/Ref./Réf.	Anmeldung Nr./Application No./Demande n°.//Patent Nr./Patent No./Brevet n°.
EPB93AC INO HC3	94915550.1 - -

Anmelder/Applicant/Demandeur//Patentinhaber/Proprietor/Titulaire
N.V. INNOGENETICS S.A.

COMMUNICATION

The European Patent Office herewith transmits

- ☐ the European search report
- ☐ the declaration under Rule 45 EPC
- ☐ the partial European search report under Rule 45 EPC
- ☒ the supplementary European search report concerning the international application under Article 157(2) EPC relating to the above-mentioned European patent application. Copies of the documents cited in the search report are enclosed.

The following specifications given by the applicant have been approved by the Search Division :

☐ Abstract ☐ Title ☐ Figure

☐ The abstract was modified by the Search Division and the definitive text is attached to this communication.

☐ The following figure will be published with the abstract, since the Search Division considers that it better characterises the invention than the one indicated by the applicant.

Figure:

☐ Additional copy(copies) of the documents cited in the European search report.

Ro

REFUND OF THE SEARCH FEE

If applicable under Article 10 Rules relating to fees, a separate communication from the Receiving Section on the refund of the search fee will be sent later.

European Patent Office

SUPPLEMENTARY EUROPEAN SEARCH REPORT

Application Number
EP 94 91 5550

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.5)
X	BIOCHEM. BIOPHYS. RES. COMMUN., vol. 183, no. 1, 1992, pages 334-342, XP002017142 S. MORI ET AL.: "A new type of hepatitis C in patients in Thailand" * figure 1 * ---	1,2, 6-11, 15-23	C12N15/51 C12Q1/68 A61K39/29 G01N33/576 C12Q1/70
X,P	BIOCHEM. BIOPHYS. RES. COMMUN., vol. 192, no. 2, 1993, pages 635-641, XP002017143 L. STUYVER ET AL.: "Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3" * figure 1 * ---	1,2, 6-11, 15-23	
X	PROC. NATL. ACAD. SCI. USA, vol. 87, 1990, pages 9254-9258, XP002017144 N. KATO ET AL.: "Molecular cloning of the human hepatitis C virus genome form Japanese patients with non-A non-B hepatitis" * figure 2 * ---	1,2, 6-11, 15-23	
X,P	WO-A-93 10239 (COMMON SERVICES AGENCY) * figure 9A * ---	1,2, 6-11, 15-23	
X	WO-A-92 19743 (CHIRON CORPORATION) * figure 2 * --- -/--	1,2, 6-11, 15-23	

TECHNICAL FIELDS SEARCHED (Int.Cl.5)
C07K
C12N
C12Q
A61K
G01N

The supplementary search report has been drawn up for the claims attached hereto.

Place of search	Date of completion of the search	Examiner
THE HAGUE	29 October 1996	Skelly, J

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C09)

European Patent Office

SUPPLEMENTARY EUROPEAN SEARCH REPORT

Application Number
EP 94 91 5550

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.5)
X,P	DDBJ database entry HPCN S5P5 Accession no. D26387, 4 February 1994, HOTTA, H ET AL.:"Subtype analysis of hepatitis C virus in Indonesia" XP002017146 *abstract*	1,2, 6-11, 15-23	
X,P	EMBL/GENBANK/DDBJ database entry HC3NL96 Accession no. X78863, 20 April 1994 VAN DOORN, J. ET AL. "Sequence analysis of hepatitis C virus genotypes 1 to 5" XP002017147 *abstract*	1,2, 5-11, 15-23	
E	WO-A-95 01442 (GOVERNMENT OF THE UNITED STATES OF AMERICA) * the whole document *	3,6-10, 12,15-23	
X	EP-A-0 532 167 (IMMUNO JAPAN INC.) *SEQ. ID. 4*	1,5-10, 14-23	
X	BIOCHEM. BIOPHYS. RES. COMMUN., vol. 170, no. 3, 1990, pages 1021-1025, XP002017145 N. ENOMOTO ET AL.: "There are two major types of hepatitis C virus in Japan" * figure 1 *	1,5-10, 14-23	

TECHNICAL FIELDS SEARCHED (Int.Cl.5)

The supplementary search report has been drawn up for the claims attached hereto.

Place of search	Date of completion of the search	Examiner
THE HAGUE	29 October 1996	Skelly, J

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C04)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 94 91 5550

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

29-10-1996

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9310239	27-05-93	AU-B- 671967	19-09-96
		AU-A- 3088792	15-06-93
		CA-A- 2123875	27-05-93
		DE-T- 610436	03-08-95
		EP-A- 0610436	17-08-94
		ES-T- 2065863	01-03-95
		FI-A- 942369	19-07-94
		JP-T- 7501442	16-02-95
		ZA-A- 9209015	17-05-93
WO-A-9219743	12-11-92	AU-B- 668355	02-05-96
		AU-A- 2155892	21-12-92
		BG-A- 98200	31-01-95
		CZ-A- 9302377	13-04-94
		EP-A- 0585398	09-03-94
		HU-A- 69609	28-09-95
		JP-T- 6508026	14-09-94
		NO-A- 934019	05-11-93
		PL-B- 169880	30-09-96
		SK-A- 123293	08-06-94
WO-A-9501442	12-01-95	US-A- 5514539	07-05-96
		AU-A- 7319194	24-01-95
EP-A-532167	17-03-93	JP-A- 6121689	06-05-94
		JP-A- 6133778	17-05-94
		CA-A- 2075611	10-02-93
		US-A- 5428145	27-06-95

European Patent Office

CLAIMS INCURRING FEES

The present European patent application comprised at the time of filing more than ten claims.

☐ All claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for all claims.

☐ Only part of the claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims and for those claims for which claims fees have been paid,

namely claims:

☐ No claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims.

LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirement of unity of invention and relates to several inventions or groups of inventions,
namely:

see sheet -B-

☒ All further search fees have been paid within the fixed time limit. The present European search report has been drawn up for all claims.

☐ Only part of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the inventions in respect of which search fees have been paid,

namely claims:

☐ None of the further search fees has been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the invention first mentioned in the claims,

namely claims:

European Patent Office

LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirement of unity of invention and relates to several inventions or groups of inventions, namely:

The separate inventions are therefore considered to be as follows:-

1. Claims 1 and 2 (partially), 6-10 (partially), 11, 15-23 (partially): Polynucleotides or amino acids corresponding to the core/E1 region of HCV subtype 3a and their uses.

2. Claims 1 and 2 (partially), 6-10 (partially), 11, 15-23 (partially): Polynucleotides or amino acids corresponding to other regions of the genome of HCV subtypes 3, 3a and 3c and their uses.

3. Claims 1 (partially), 3, 6-10 (partially), 13, 15-23 (partially): Polynucleotides or amino acids corresponding to various regions of the genome of HCV subtype 5 and their uses.

4. Claims 1 (partially), 4, 6-10 (partially), 12, 15-23 partially: Polynucleotides or amino acids corresponding to various regions of the genome of HCV subtype 4 and their uses.

5. Claims 1, 5, 6-10 (partially), 14, 15-23 (partially): Polynucleotides or amino acids corresponding to various regions of the genome of HCV subtype 2d and their uses.

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

United States Patent and Trademark Office
(Box PCT)
Washington D.C. 20231
United States of America

in its capacity as elected Office

Date of mailing (day/month/year) 13 June 1996 (13.06.96)	
International application No. PCT/EP95/04155	**Applicant's or agent's file reference** PCT.95.46.HCV
International filing date (day/month/year) 23 October 1995 (23.10.95)	**Priority date** (day/month/year) 21 October 1994 (21.10.94)
Applicant MAERTENS, Geert et al	

1. The designated Office is hereby notified of its election made:

   [X] in the demand filed with the International Preliminary Examining Authority on:

   17 May 1996 (17.05.96)

   [ ] in a notice effecting later election filed with the International Bureau on:

   _____

2. The election [X] was

   [ ] was not

   made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland**	Authorized officer M. Abidine
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 730.91.11

Copy for the Elected Office (EO/US)

PCT/EP95/04155

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92bis.1 and Administrative Instructions, Section 422)

To:

INNOGENETICS N.V.
Mrs. Ann De Clercq
Industriepark Zwijnaarde 7
Box 4
B-9052 Zwijnaarde
BELGIQUE

Date of mailing (day/month/year)
25 July 1996 (25.07.96)

Applicant's or agent's file reference
PCT.95.46.HCV

IMPORTANT NOTIFICATION

International application No.
PCT/EP95/04155

International filing date (day/month/year)
23 October 1995 (23.10.95)

1. The following indications appeared on record concerning:

[X] the applicant [X] the inventor [ ] the agent [ ] the common representative

Name and Address
STUYVER, Lieven
Holestraat 8
B-2400 Mol
Belgium

State of Nationality

State of Residence

Telephone No.

Facsimile No.

Teleprinter No.

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

[ ] the person [ ] the name [X] the address [ ] the nationality [ ] the residence

Name and Address
STUYVER, Lieven
Holestraat 8
B-9552 Herzele
Belgium

State of Nationality

State of Residence

Telephone No.

Facsimile No.

Teleprinter No.

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

[X] the receiving Office
[ ] the International Searching Authority
[X] the International Preliminary Examining Authority
[ ] the designated Offices concerned
[X] the elected Offices concerned
[ ] other:

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer
M. Abidine

Telephone No.: (41-22) 730.91.11

PATENT COOPERATION TREAT

From the INTERNATIONAL BUREAU

PCT

COMMUNICATION OF
INTERNATIONAL APPLICATIONS

(PCT Article 20)

To:

United States Patent and Trademark
Office
(Box PCT)
Crystal Plaza 2
Washington, DC 20231
ETATS-UNIS D'AMERIQUE

in its capacity as designated Office

Date of mailing:
15 August 1996 (15.08.96)

The International Bureau transmits herewith copies of the international applications having the following international application numbers and international publication numbers:

International application no.:	International publication no.:
PCT/EP95/04155	WO96/13590

CORRECTED VERSION
VERSION CORRIGEE

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer:

J. Zahra

Telephone No.: (41-22) 730.91.11

TENT COOPERATION TREA

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
PCT.95.46.HCV		
International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/EP 95/ 04155	23/10/1995	21/10/1994
International Patent Classification (IPC) or national classification and IPC		
C12N15/40		
Applicant		
INNOGENETICS N.V. et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 7 sheets, including this cover sheet.

☐ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consists of a total of _____ sheets.

3. This report contains indications and corresponding pages relating to the following items:

- I ☒ Basis of the report
- II ☐ Priority
- III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV ☐ Lack of unity of invention
- V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☐ Certain documents cited
- VII ☐ Certain defects in the international application
- VIII ☒ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
17/05/1996	18. 11. 96
Name and mailing addr ss of the IPEA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk - Netherlands Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	M. Cupido
	Telephone No.

I. Basis of the rep rt

1. This report has been drawn up on the basis of *(Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.)*

X the international application as originally filed

O the description, pages , as originally filed

pages , filed with the demand

pages , filed with the letter of

O the claims, Nos. , as originally filed

Nos. , as amended under Article 19

Nos. , filed with the demand

Nos. , filed with the letter of

O the drawings, sheets / fig. , as originally filed

sheets / fig. , filed with the demand

sheets / fig. , filed with the letter of

2. The amendments have resulted in the cancellation of:

O the description, pages:

O the claims, Nos.

O the drawings, sheets / fig.

3. O This report has been established as if (some of) the amendments had not been made, since they have been considered t go beyond the disclosure as filed (Rule 70.2 (c)).

4. Additional observations, if necessary:

III. **Non-establishment of opinion with regard to novelty, inventive step and industrial applicability**

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

o the entire international application,

x claims Nos. 37, 38, 50, 51, 61 and 62

because:

x the said international application, or the said claims relate to the following Nos. subject matter which does not require an international preliminary examination *(specify)*:

claims 37, 38, 50, 51, 61 and 62 are directed to a method of treatment of the human or animal body, see Article 34(4)(a)(i) and Rule 67.1(iv).

For the assessment of the present claims 37, 38, 50, 51, 61 and 62 on the question whether they are industrially applicable, no unified criteria exist in the PCT. The patentability can also be dependent upon the formulation of the claims. The EPO, for example, does not recognize as industrially applicable the subject-matter of claims to the use of a compound in medical treatment, but will allow, however, claims to a known compound for first use in medical treatment and the use of such a compound for the manufacture of a medicament for a new medical treatment.

o the description, claims or drawings *(indicate particular elements below)* or Nos. said claims are so unclear that no meaningful opinion could be formed *(specify)*:

o the claims, or said claims are so inadequately supported by the description Nos. no meaningful opinion could be formed.

x no international search report has been established for said claims Nos. 1-62 (partly and as far as applicable)

The IPEA concurs with the objection put forward by the ISA that the application lacks unity of invention within the meaning of Rule 13(1), for the following reasons:

1. Nucleotide sequences from many HCV subtypes have been disclosed in the prior art, see for example Proc.Natl.Acad.Sci.USA 91:10134-10138, (D1) figure 1. In the light of this document, the problem addressed by the present application is the provision of nucleotide sequences from further HCV subtypes or types to be used in diagnostic tests. The solutions provided and claimed are

sequences obtained from 14 different HCV subtypes or types, the encoded polypeptides and antibodies thereto.

2. In view of the fact that polynucleotide sequences from many HCV subtypes, the encoded polypeptides and antibodies specific thereto have been disclosed in the prior art, due to the essential difference of the HCV subtypes, and due to the fact that no other technical feature can be distinguished which defines the contribution each of the claimed inventions makes over the prior art, there is no single inventive concept underlying the plurality of the claimed inventions of the present application in the sense of Rule 13.1. Consequently the present application lacks unity of invention and the different inventions, not belonging to a common inventive concept, are the following:

1 Claims 1-62 (partly): Nucleic acids from an HCV subtype 1d sequence, use thereof for diagnostic purposes, peptides encoded and corresponding antibodies and kits containing primers or probes derived from HCV subtype 1d.

2-14 Claims 1-62 (partly): As above, from HCV subtypes 1e, 1f, 2e, 2f, 2g, 2h, 2i, 4k, 41 or HCV types 7, 8, 9 or 10.

The international search was restricted to the first invention, relating to nucleic acids from an HCV subtype 1d sequence, use thereof for diagnostic purposes, peptides encoded and corresponding antibodies and kits containing primers or probes derived from HCV subtype 1d. As a consequence, the preliminary examination can only refer to the same subject-matter.

V. **Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citati ns and explanations supporting such statement**

1. Statement

Novelty	Claims	6, 25-62	YES
	Claims	1-5, 7-24	NO
Inventive Step	Claims		YES
	Claims	1-62	NO
Industrial Applicability	Claims	1-62	YES
	Claims		NO

2. Citations and Explanations

I Documents

The following documents have been considered for the purposes of this report:

D1: Proc.Natl.Acad.Sci.USA **91**:10134-10138 (L.Stuyver et al.; 1994)
D2: J.Hepatology **21**:121-129 (L-J Van Doorn et al.; 1994)
D3: Gene **114**:245-250 (K.Liu et al.; 1992)

II Novelty

1. D2 states on page 122 that homologies between isolates belonging to the same subtypes usually exceed 90%. As only partial sequences are available from the present description, it is assumed that HCV strains showing more than 90% homology with the nucleotide sequences encoding the peptide fragments from BNL1 and BNL2 provided are within the claimed subject-matter.

2. The sequence in D3 contains 33 contiguous nucleotides (numbers 577 to 610) identical to the claimed HCV subtype 1d sequence. The corresponding amino acid sequence EVRNASGVYHV is given in figure 2, which contains a contiguous sequence of 11 amino-acids corresponding to HCV 1d SEQ ID NO 120. Hence, the sequence of the HCV-PRC1 genome represented in figure 2 of D3 is regarded to be of the HCV subtype 1d and the subject-matter of claims 1-5, 7-24 insofar it relates to the HCV subtype 1d is regarded to lack novelty within the meaning of Article 33(2).

III Inventive step

D1 is regarded as the closest prior art with respect to the question whether the claimed subject-matter involves an inventive step. The problem underlying the present invention in view of D1 is to provide further HCV sequences, useful to diagnose the presence of HCV type 1. No data have been provided in the description, demonstrating that the claimed sequences , or the encoded peptides or antibodies raised upon immunisation with such peptides, are a useful as a diagnostic tool. In the absence of such data the presence of an inventive step within the meaning of Article 33(3), cannot be acknowledged.

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by th description, are made:

1. The term "...heretofore unidentified..." HCV type or subtype used in claim 1 is not a technical feature and cannot be accepted to distinguish the claimed subject-matter from the prior art. The same remark is made in the context of claims 6, 22, 23 and 41which refer to "...known..." HCV types. Hence claims 1, 22, 23 and 41 are not clear and concise as required by Article 6.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/40, C07K 14/18, C12Q 1/70, C07K 16/10, G01N 33/569, A61K 39/29

A3

(11) International Publication Number: **WO 96/13590**

(43) International Publication Date: 9 May 1996 (09.05.96)

(21) International Application Number: PCT/EP95/04155

(22) International Filing Date: 23 October 1995 (23.10.95)

(30) Priority Data:
94870166.9 21 October 1994 (21.10.94) EP
(34) Countries for which the regional or international application was filed: AT et al.
95870076.7 28 June 1995 (28.06.95) EP
(34) Countries for which the regional or international application was filed: AT et al.

(71) Applicant *(for all designated States except US)*: INNOGENETICS N.V. [BE/BE]; Industriepark Zwijnaarde 7, Box 4, B-9052 Zwijnaarde (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MAERTENS, Geert [BE/BE]; Zilversparrenstraat 64, B-8310 Brugge (BE). STUYVER, Lieven [BE/BE]; Holestraat 8, B-2400 Mol (BE).

(74) Common Representative: INNOGENETICS N.V.; Mrs. Ann De Clercq, Industriepark Zwijnaarde 7, Box 4, B-9052 Zwijnaarde (BE).

(81) Designated States: AL, AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, LS, MW, SD, SZ, UG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search r port:
15 August 1996 (15.08.96)

(54) Title: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC AGENTS

(57) Abstract

The present invention relates to new genomic nucleotide sequences and amino acid sequences corresponding to the coding region of these genomes. The invention relates to new HCV types and subtypes sequences which are different from the known HCV types and subtypes sequences. More particularly, the present invention relates to new HCV type 7 sequences, new HCV type 9 sequences, new HCV type 10 and new HCV type 11 sequences. Also, the present invention relates to new HCV type 1 sequences of subtypes 1d, 1e, 1f and 1g; new HCV type 2 sequences of subtypes 2e, 2f, 2g, 2h, 2i, 2k and 2l; new HCV type 3 sequences of subtype 3g, new HCV type 4 sequences of subtypes 4k, 4l and 4m; a process for preparing them, and their use for diagnosis, prophylaxis and therapy. More particularly, the present invention provides new type-specific sequences of the Core, the E1 and the NS5 regions of new HCV types 7, 9, 10 and 11, as well as of new variants (subtypes) of HCV types 1, 2, 3 and 4. These new HCV sequences are useful to diagnose the presence of HCV type 1, and/or type 2, and/or type 3, and/or type 4, and/or type 7, and/or type 9, and/or type 10, and/or type 11 genotypes or serotypes in a biological sample. Moreover, the availability of these new type-specific sequences can increase the overall sensitivity of HCV detection and should also prove to be useful for prophylactic and therapeutic purposes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SI	Slovenia
SK	Slovakia
SN	Senegal
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/40, C07K 14/18, C12Q 1/70, C07K 16/10, G01N 33/576 — A2

(11) International Publication Number: **WO 96/13590**

(43) International Publication Date: 9 May 1996 (09.05.96)

(21) International Application Number: PCT/EP95/04155

(22) International Filing Date: 23 October 1995 (23.10.95)

(30) Priority Data:
94870166.9 21 October 1994 (21.10.94) EP
(34) Countries for which the regional or international application was filed: AT et al.
95870076.7 28 June 1995 (28.06.95) EP
(34) Countries for which the regional or international application was filed: AT et al.

(71) Applicant *(for all designated States except US)*: INNOGENETICS N.V. [BE/BE]; Industriepark Zwijnaarde 7, Box 4, B-9052 Zwijnaarde (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MAERTENS, Geert [BE/BE]; Zilversparrenstraat 64, B-8310 Brugge (BE). STUYVER, Lieven [BE/BE]; Holestraat 8, B-2400 Mol (BE).

(74) Common Representative: INNOGENETICS N.V.; Mrs. Ann De Clercq, Industriepark Zwijnaarde 7, Box 4, B-9052 Zwijnaarde (BE).

(81) Designated States: AL, AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, LS, MW, SD, SZ, UG).

Published
Without international search report and to be republished upon receipt of that report.
With a request for rectification under Rule 91.1(f).

(54) Title: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC AGENTS

(57) Abstract

The present invention relates to new genomic nucleotide sequences and amino acid sequences corresponding to the coding region of these genomes. The invention relates to new HCV types and subtypes sequences which are different from the known HCV types and subtypes sequences. More particularly, the present invention relates to new HCV type 7 sequences, new HCV type 9 sequences, new HCV type 10 and new HCV type 11 sequences. Also, the present invention relates to new HCV type 1 sequences of subtypes 1d, 1e, 1f and 1g; new HCV type 2 sequences of subtypes 2e, 2f, 2g, 2h, 2i, 2k and 2l; new HCV type 3 sequences of subtype 3g, new HCV type 4 sequences of subtypes 4k, 4l and 4m; a process for preparing them; and their use for diagnosis, prophylaxis and therapy. More particularly, the present invention provides new type-specific sequences of the Core, the E1 and the NS5 regions of new HCV types 7, 9, 10 and 11, as well as of new variants (subtypes) of HCV types 1, 2, 3 and 4. These new HCV sequences are useful to diagnose the presence of HCV type 1, and/or type 2, and/or type 3, and/or type 4, and/or type 7, and/or type 9, and/or type 10, and/or type 11 genotypes or serotypes in a biological sample. Moreover, the availability of these new type-specific sequences can increase the overall sensitivity of HCV detection and should also prove to be useful for prophylactic and therapeutic purposes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

3c, 3d, 3e, 3f, 4a, 4b, 4c, 4d, 4e, 4f, 4g, 4h, 4i, 4j, 5a or 6a, with said HCV subtypes being classified as in Table 3 by comparison of a part of the NS5 gene nucleotide sequence spanning positions 7932 to 8271, with said amino acid numbering being shown in Table 1, and with said polynucleic acid containing at least one nucleotide differing from said known HCV nucleotide sequences, or the complement thereof. The sequence of known HCV isolates may be found in any nucleotide sequence database known in the art (such as for instance the EMBL database).

The present invention thus also relates to a polynucleic acid having a nucleotide sequence which is unique to at least one of HCV subtypes 1d, 1e, 1f, 1g, 2e, 2f, 2g, 2h, 2i, 2k, 2l, 3g, 4k, 4l, 4m, 7a, 7c or 7d, with said HCV subtypes being classified as defined above.

The present invention thus also relates to a polynucleic acid having a nucleotide sequence which is unique to at least one of HCV types 9, 10 or 11, with said HCV types being classified as defined above.

It is to be noted that the nucleotide(s) difference in the polynucleic acids of the invention may involve an amino acid difference in the corresponding amino acid sequences encoded by said polynucleic acids. A composition according to the present invention may contain only polynucleic acid sequences or polynucleic acid sequences mixed with any excipient known in the art of diagnosis, prophylaxis or therapy.

According to a preferred embodiment, the present invention relates to a polynucleic acid encoding an HCV polyprotein comprising in its amino acid sequence at least one of the following amino acid residues:

I15, C38, V44, A49, Q43, P49, Q55, A58, S60 or D60, E68 or V68, H70, A71 or Q71 or N71, D72, H81, H101, D106, S110, L130, I134, E135, L140, S148, T150 or E150, Q153, F155, D157, G160, E165, I169, F181, L186, T190, T192 or I192 or H192, I193, A195, S196, R197 or N197 or K197, Q199 or D199 or H199 or N199, F200 or T200, A208, I213, M216 or S216, N217 or S217 or G217 or K217, T218, I219, A222, Y223, I230, W231 or L231, S232 or H232 or A232, Q233, E235 or L235, F236 or T236, F237, L240 or M240, A242, N244, N249, I250 or K250 or R250, A252 or C252, A254, I255 or V255, D256 or M256, E257, E260 or K260, R261, V268, S272 or R272, I285, G290 or F290, A291, A293 or L293

or W293, T294 or A294, S295 or H295, K296 or E296, Y297 or M297, I299 or Y299, I300, S301, P316, S2646, A2648, G2649, A2650, V2652, Q2653, H2656 or L2656, D2657, F2659, K2663 or Q2663, A2667 or V1667, D2677, L2681, M2686 or Q2686 or E2686, A2692 or K2692, H2697, I2707, L2708 or Y2708, A2709, A2719 or M2719, F2727, T2728 or D2728, E2729, F2730 or Y2730, I2741, I2745, V2746 or E2746 or L2746 or K2746, A2748, S2749 or P2749, R2750, E2751, D2752 or N2752 or S2752 or T2752 or V2752 or I2752 or Q2752, S2753 or D2753 or G2753, D2754, A2755, L2756 or Q2756, R2757,

with said notation being composed of a letter representing the amino acid residue by its one-letter code, and a number representing the amino acid numbering according to Kato et al. (1980), as shown in Table 1,

or a part of said polynucleic acid which is unique to at least one of the HCV subtypes or types as defined in Table 5, and which contains at least one nucleotide differing from known HCV nucleotide sequences, or the complement thereof.

Each of the above-mentioned residues can be found in Figures 2, 4 or 6 showing the new amino acid sequences of the present invention aligned with known sequences of other types or subtypes of HCV for the Core/E1 region.

According to another preferred embodiment, the present invention relates to a polynucleic acid encoding a HCV polyprotein comprising in its amino acid sequence at least one amino acid sequence chosen from the following list:

ARQSDGRSWAQ or ARRSEGRSWAQ as for subtype 1d (SEQ ID NO 107 and 108)

ERRPEGRSWAQ as for subtype 1e (SEQ ID NO 109)

ARRPEGRSWAQ as for subtype 1f (SEQ ID NO 110)

DRRTTGKSWGR as for subtype 2k (SEQ ID NO 111)

DRRATGRSWGR as for subtype 2e (SEQ ID NO 112)

DRRATGKSWGR as for subtype 2f (SEQ ID NO 113)

VRQPTGRSWGQ as for type 9 (SEQ ID NO 114)

VRHQTGRTWAQ as for subtype 7a and 7c (SEQ ID NO 115)

VRQNQGRTWAQ as for subtype 7d (SEQ ID NO 116)

ARRTEGRSWAQ as for type 10 (SEQ ID NO 117)

VRRTTGRXXXX or VRRTTGRTWAQ as for type 11 and 119) (SEQ ID NO 118

HEVRNASGVYHV or HEVRNASGVYHL as for subtype 1d (SEQ ID NO 120 and 121)

YEVHSTTDGYHV as for subtype 1f (SEQ ID NO 122)

VEVKNTSQAYMA as for subtype 2e (SEQ ID NO 123)

IQVKNNSHFYMA as for subtype 2f (SEQ ID NO 124)

VQVKNTSTMYMA as for subtype 2g (SEQ ID NO 125)

VQVKNTSHSYMV as for subtype 2h (SEQ ID NO 126)

VQVANRSGSYMV as for subtype 2i (SEQ ID NO 127)

VEIKNTXNTYVL or VEIKNTSNTYVL as for subtype 2k (SEQ ID NO 128 and 129)

INYRNVSGIYYV or INYRNTSGIYHV or INYHNTSGIYHI or TNYRNVSGIYHV as for subtype 4k (SEQ ID NO 130, 131, 132 or 133)

QHYRNVSGIYHV as for subtype 4l (SEQ ID NO 134)

IQVKNASGIYHL as for type 9 (SEQ ID NO 135)

AHYTNKSGLYHL as for subtype 7c (SEQ ID NO 136)

LNYANKSGLYHL as for subtype 7d (SEQ ID NO 137)

LEYRNASGLYMV as for type 10 (SEQ ID NO 138)

IYEMDGMIMHY or IYEMSGMILHA as for subtype 1d (SEQ ID NO 139 and 140)

VYEAKDIILHT as for subtype 1f (SEQ ID NO 141)

VWQLXDAVLHV as for subtype 2e (SEQ ID NO 142)

VWQLRDAVLHV as for subtype 2f (SEQ ID NO 143)

IWQMQGAVLHV as for subtype 2g (SEQ ID NO 144)

VWQLKDAVLHV as for subtype 2h (SEQ ID NO 145)

VWQLEEAVLHV as for subtype 2i (SEQ ID NO 146)

TWQLXXAVLHV as for subtype 2k (SEQ ID NO 147)

VYEADHHILHL or VYEADHHILAL or VFEADHHILHL as for subtupe 4k (SEQ ID NO 148, 149 and 150)

VYESDHHILHL as for subtype 4l (SEQ ID NO 151)

VFEAETMILHL as for type 9 (SEQ ID NO 152)

VYEAETLILHL as for subtype 7c (SEQ ID NO

153)

VYEANGMILHL as for subtype 7d (SEQ ID NO 154)

VYEAGDIILHL as for type 10 (SEQ ID NO 155)

VREDNHLRCWMAL or VRENNSSRCWMAL as for subtype 1d (SEQ ID NO 156 and 157)

IREGNISRCWVPL as for subtype 1f (SEQ ID NO 158)

ENSSGRFHCWIPI as for subtype 2e (SEQ ID NO 159)

ERSGNRTFCWTAV as for subtype 2f (SEQ ID NO 160)

ELQGNKSRCWIPV as for subtype 2g (SEQ ID NO 162)

ERHQNQSRCWIPV as for subtype 2h (SEQ ID NO 163)

EWKDNTSRCWIPV as for subtype 2i (SEQ ID NO 164)

EREGNSSRCWIPV as for subtype 2k (SEQ ID NO 165)

VREGNQSRCWVAL or VRTGNQSRCWVAL or VRVGNQSSCWVAL or VRVGNQSRCWVAL or VKEGNHSRCWVAL as for subtype 4k (SEQ ID NO 166, 167, 168 or 169)

VKTGNTSRCWVAL as for subtype 4l (SEQ ID NO 170)

IKAGNESRCWLPV as for type 9 (SEQ ID NO 171)

VKEGNQSRCWVQA as for subtype 7c (SEQ ID NO 172)

VKXXNLTKCWLSA as for subtype 7d (SEQ ID NO 173)

VRSGNTSRCWIPV as for type 10 (SEQ ID NO 174)

VKNASVPTAA or VKDANVPTAA as for subtype 1d (SEQ ID NO 175 and 176)

ARIANAPIDE as for subtype 1f (SEQ ID NO 177)

VSKPGALTKG as for subtype 2e (SEQ ID NO 178)

VSRPGALTRG as for subtype 2f (SEQ ID NO 179)

VNQPGALTRG as for subtype 2g (SEQ ID NO 180)

VSQPGALTRG as for subtype 2h (SEQ ID NO 181)

VSQPGALTKG as for subtype 2i (SEQ ID NO 182)

VSRPGALTEG as for subtype 2k (SEQ ID NO 183)

APYIGAPLES or APYTAAPLES as for subtype 4k (SEQ ID NO 184 and 185)

APILSAPLMS as for subtype 4l (SEQ ID NO 186)

VPNSSVPIHG as for type 9 (SEQ ID NO 187)

VPNASTPVTG as for subtype 7c (SEQ ID NO 188)

VQNASVSIRG as for subtype 7d (SEQ ID NO 189)
VKSPCAATAS as for type 10 (SEQ ID NO 190)
SPRMHHTTQE or SPRLYHTTQE as for subtype 1d (SEQ ID NO 191 and 192)
TSRRHWTVQD as for subtype 1f (SEQ ID NO 193)
APKRHYFVQE as for subtype 2e (SEQ ID NO 194)
SPQYHTFVQE as for subtype 2f (SEQ ID NO 195)
SPQHHNFSQD as for subtype 2g (SEQ ID NO 196)
SPQHHIFVQD as for subtype 2h (SEQ ID NO 197)
SPEHHHFVQD as for subtype 2k (SEQ ID NO 198)
RPRRHWTTQD or RPRRHWTAQD or QPRRHWTTQD or RPRRHWTTQE as for subtype 4k (SEQ ID NO 199, 200, 201 or 202)
QPRRHWTVQD as for subtype 4l (SEQ ID NO 203)
RPKYHQVTQD as for type 9 (SEQ ID NO 204)
RPRMHQVVQE as for subtype 7c (SEQ ID NO 205)
RPRMYEIAQD as for subtype 7d (SEQ ID NO 206)
RHRQHWTVQD as for type 10 (SEQ ID NO 207)

or a part of said polynucleic acid which is unique to at least one of the HCV subtypes or types as defined Table 5, and which contains at least one nucleotide differing from known HCV nucleotide sequences, or the complement thereof.

Using the 5' non-coding LiPA system (Stuyver et al., 1993) and a new core LiPA system including multiple probes for subtypes 1a, 1b, 1c, 2a, 2b or 2c derived from the core region (Stuyver et al., 1995), samples from the Benelux, Cameroon, France and Vietnam were selected because of their aberrant reactivities (isolates CAM1078, FR2, FR1, VN4, VN12, VN13, NE98). Some samples were, together with many other samples, sequenced as a control for typing. Sequencing results, however, indicated the discovery of new subtypes (isolates BNL1, BNL2, BNL3, FR4, BNL4, BNL5, BNL6, BNL7, BNL8, BNL9, BNL10, BNL11 and BNL12). Nucleotide sequences in the core and E1 regions which have not yet been reported before, were analyzed in the frame of the invention. Genomic sequences of subtype 1d, 1e, 1f, 1g 2e, 2f, 2g, 2h, 2i, 2k, 2l, 3g, 4k, 4l, 4m, 7a, 7c, 7d and types 9, 10 and 11 isolates are reported for the first time in the present invention. The NS5B region was also analyzed.

The term "polynucleic acid" refers to a single- stranded or double-stranded

nucleic acid sequence which may contain at least 5 contiguous nucleotides in common with the complete nucleotide sequence (e.g. at least 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 75 or more contiguous nucleotides). A polynucleic acid which is up till about 100 nucleotides in length is often also referred to as an oligonucleotide. A polynucleic acid may consist of deoxyribonucleotides or ribonucleotides, nucleotide analogues or modified nucleotides, or may have been adapted for therapeutic purposes. A polynucleic acid may also comprise a double stranded cDNA clone which can be used for cloning purposes, or for *in vivo* therapy, or prophylaxis.

The oligonucleotides according to the present invention, used as primers or probes may also contain or consist of nucleotide analogous such as phosphorothioates (Matsukura et al., 1987), alkylphosphoriates (Miller et al., 1979) or peptide nucleic acids (Nielsen et al., 1991; Nielsen et al., 1993) or may contain interculating agents (Asseline et al., 1984).

As most other variations or modifications introduced into the original DNA sequences of the invention these variations will neccissitate adaptions with respect to the conditions under which the oligonucleotide should be used to obtain the required specificty and sensitivity. However the eventual results will be essentially the same as those obtained with the unmodified oligonucleotides.

The introduction of these modifications may be advantageous in order to positivily influence characteristics such as hybridization kinetics, reversibility of the hybrid-formation, biological stability of the oligonucleotide molecules, etc.

The polynucleic acids of the invention may be comprised in a composition of any kind. Said composition may be for diagnostic, therapeutic or prophylactic use.

The expression "sequences which are unique to an HCV type or subtype" refers to sequences which are not shared by any other type or subtype of HCV, and can thus be used to uniquely detect that HCV type or subtype. Sequence variability is demonstrated in the present invention between the newly found HCV types and subtypes (see Table 5) and the known HCV types and subtypes (see Table 3), and it is therefore from these regions of sequence variability in particular that type- or subtypes-specific polynucleic acids, oligonucleotides, polypeptides and peptides may be obtained. The term type- or subtypes-specific refers to the fact that a sequence is unique to that HCV type or subtype involved.

The expression "nucleotides corresponding to" refers to nucleotides which are homologous or complementary to an indicated nucleotide sequence or region within a specific HCV sequence.

The term "coding region" corresponds to the region of the HCV genome that encodes the HCV polyprotein. In fact, it comprises the complete genome with the exception of the 5' untranslated region and 3' untranslated region.

The term "HCV polyprotein" refers to the HCV polyprotein of the HCV-J isolate (Kato et al., 1990). The adenine residue at position 330 (Kato et al., 1990) is the first residue of the ATG codon that initiates the long HCV polyprotein of 3010 amino acids in HCV-J and other type 1b isolates, and of 3011 amino acids in HCV-1 and other type 1a isolates, and of 3033 amino acids in type 2 isolates HC-J6 and HC-J8 (Okamoto et al., 1992).

This adenine is designated as position 1 at the nucleic acid level, and this methionine is designated as position 1 at the amino acid level, in the present invention. As type 1a isolates contain 1 extra amino acid in the NS5A region, coding sequences of type 1a and 1b have identical numbering in the Core, E1, NS3, and NS4 region, but will differ in the NS5B region as indicated in Table 1. Type 2 isolates have 4 extra amino acids in the E2 region, and 17 or 18 extra amino acids in the NS5 region compared to type 1 isolates, and will differ in numbering from type 1 isolates in the NS3/4 region and NS5b regions as indicated in Table 1. Similar insertions compared with type 1 (but of a different size) can also be observed in type 3a sequences which affect the numbering of type 3a amino acids accordingly. Other insertions or deletions may be readily observed in type1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11 sequences after alignment withknown HCV sequences.

TABLE 1

	Region	Positions described in the present invention*	Positions described for HCV-J (Kato et al., 1990)	Positions described for HCV-1 (Choo et al., 1991)	Positions described for HC-J6, HC-J8 (Okamoto et al., 1992)

Nucleotides	NS5B	8023/8235 7932/8271	8352/8564 8261/8600	8026/8238 7935/8274	8433/8645 8342/8681
		coding region of present invention	330/9359	1/9033	342/9439
Amino Acids	NS5B	2675/2745 2645/2757	2675/2745 2645/2757	2676/2746 2646/2758	2698/2768 2668/2780

Table 1: Comparison of the HCV nucleotide and amino acid numbering system used in the present invention (*) with the numbering used for other prototype isolates. For example, 8352/8564 indicates the region designated by the numbering from nucleotide 8352 to nucleotide 8564 as described by Kato et al. (1990). Since the numbering system of the present invention starts at the polyprotein initiation site, the 329 nucleotides of the 5' untranslated region described by Kato et al. (1990) have to be substracted, and the corresponding region is numbered from nucleotide 8023 ('8352-329') to 8235 ('8564-329').

The term "genotype" as used in the present invention refers to both types and/or subtypes.

The term "HCV type" corresponds to a group of HCV isolates of which the complete genome shows more than 73% preferably more than 74% homology at the nucleic acid level, or of which the NS5 region between nucleotide positions 7932 and 8271 shows more than 75.4% homology at the nucleic acid level, or of which the complete HCV polyprotein shows more than 78% homology at the amino acid level, or of which the NS5 region between amino acids at positions 2645 and 2757 shows more than 80% homology at the amino acid level, to polyproteins of the other isolates of the group, with said numbering beginning at the first ATG codon or first methionine of the long HCV polyprotein of the HCV-J isolate (Kato et al., 1990). Isolates belonging to different types of HCV exhibit homologies, over the complete genome, of less than 74%, preferably less than 73%, at the nucleic acid level and less than 78% at the amino acid level. Isolates belonging to the same type usually

show homologies of about 90 to 99% at the nucleic acid level and 95 to 96% at the amino acid level when belonging to the same subtype, and those belonging to the same type but different subtypes preferably show homologies of about 76% to 82% (more particularly of about 77% to 80%) at the nucleic acid level and 85-86% at the amino acid level.

More preferably the definition of HCV types is concluded from the classification of HCV isolates according to their nucleotide distances calculated as detailed below:

(1) based on phylogenetic analysis of nucleic acid sequences in the NS5B region between nucleotides 7935 and 8274 (Choo et al., 1991) or 8261 and 8600 (Kato et al., 1990) or 8342 and 8681 (Okamoto et al., 1991), isolates belonging to the same HCV type show nucleotide distances of less than 0.34, usually less than 0.33, and more usually of less than 0.32, and isolates belonging to the same subtype show nucleotide distances of less than 0.135, usually of less than 0.13, and more usually of less than 0.125, usually ranging between 0.0003 and 0.1151, and consequently isolates belonging to the same type but different subtypes show nucleotide distances ranging from 0.135 to 0.34, usually ranging from 0.1384 to 0.2977, and more usually ranging from 0.15 to 0.32, and isolates belonging to different HCV types show nucleotide distances greater than 0.34, usually greater that 0.35, and more usually of greater than 0.358, more usually ranging from 0.3581 to 0.6670.

(2) based on phylogenetic analysis of nucleic acid sequences in the core/E1 region between nucleotides 378 and 957, isolates belonging to the same HCV type show nucleotide distances of less than 0.38, usually of less than 0.37, and more usually of less than 0.364, and isolates belonging to the same subtype show nucleotide distances of less than 0.17, usually of less than 0.16, and more usually of less than 0.15, more usually less than 0.135, more usually less than 0.134, and consequently isolates belonging to the same type but different subtypes show nucleotide distances ranging from 0.15 to 0.38, usually ranging from 0.16 to 0.37, and more usually ranging from 0.17 to 0.36, more usually ranging from 0.133 to 0.379, and isolates belonging to different HCV types show nucleotide distances greater than 0.34, 0.35, 0.36, usually more than 0.365, and more usually of greater than 0.37,

Table 2 : Molecular evolutionary distances

Region	Core/E1 579 bp	E1 384 bp	NS5B 340 bp	NS5B 222 bp
Isolates*	0.0017 - 0.1347 (0.0750 ± 0.0245)	0.0026 - 0.2031 (0.0969 ± 0.0289)	0.0003 - 0.1151 (0.0637 ± 0.0229)	0.000 - 0.1323 (0.0607 ± 0.0205)
Subtypes*	0.1330 - 0.3794 (0.2786 ± 0.0363)	0.1645 - 0.4869 (0.3761 ± 0.0433)	0.1384 - 0.2977 (0.2219 ± 0.0341)	0.117 - 0.3538 (0.2391 ± 0.0399)
Types*	0.3479 - 0.6306 (0.4703 ± 0.0525)	0.4309 - 0.9561 (0.6308 ± 0.0928)	0.3581 - 0.6670 (0.4994 ± 0.0495)	0.3457 - 0.7471 (0.5295 ± 0.0627)

Table 2 Figures created by the PHYLIP program DNADIST are expressed as minimum to maximum (average ± standard deviation). Phylogenetic distances for isolates belonging to the same subtype ('isolates'), to different subtypes of the same type ('subtypes'), and to different types ('types') are given.

In a comparative phylogenetic analysis of available sequences, ranges of molecular evolutionary distances for different regions of the genome were calculated, based on 19,781 pairwise comparisons by means of the DNADIST program of the phylogeny inference package PHYLIP version 3.5c (Felsenstein, 1993). The results are shown in Table 2 and indicate that although the majority of distances obtained in each region fit with classification of a certain isolate, only the ranges obtained in the 340bp NS5B-region are non-overlapping and therefore conclusive. However, as was performed in the present invention, it is preferable to obtain sequence information from at least 2 regions before final classification of a given isolate.

Designation of a number to the different types of HCV and HCV nomenclature is based on chronological discovery of the different types. The numbering system used in the present invention might still fluctuate according to international conventions or guidelines. For example, "type 4" might be changed into "type 5" or "type 6". Also the arbitrarily chosen border distances between types and subtypes and isolates may still be subject to change according to international guidelines or

conventions. Therefore types 7a, 8a, 8b, 9a may for example be designated 6b, 6c, 6d, and 6d in the future; and type 10a which shows relatedness with genotype 3 may be denoted 3g instead of 10a.

The term "subtype" corresponds to a group of HCV isolates of which the complete polyprotein shows a homology of more than 90% both at the nucleic acid and amino acid levels, or of which the NS5 region between nucleotide positions 7932 and 8271 shows a homology of more than 90% at the nucleic acid level to the corresponding parts of the genomes of the other isolates of the same group, with said numbering beginning with the adenine residue of the initiation codon of the HCV polyprotein. Isolates belonging to the same type but different subtypes of HCV show homologies of more than 74% at the nucleic acid level and of more than 78% at the amino acid level.

It is to be understood that extremely variable regions such as the E1, E2 and NS4 regions will exhibit lower homologies than the average homology of the complete genome of the polyprotein.

Using these criteria, HCV isolates can be classified into at least 11 types. Several subtypes can clearly be distinguished in types 1, 2, 3, 4 and 7 : 1a, 1b, 1c, 1d, 1e, 1f, 1g, 2a, 2b, 2c, 2d, 2e, 2f, 2g, 2h, 2i, 2k, 2l, 3a, 3b, 3c, 3d, 3f, 3g, 4a, 4b, 4c, 4d, 4e, 4f, 4g, 4h, 4i, 4j, 4k, 4l, 4m, 7a, 7c, and 7d based on homologies of the 5' UR and coding regions. An overview of most of the reported isolates and their proposed classification according to the typing system of the present invention as well as other proposed classifications is presented in Table 3.

Table 3

HCV CLASSIFICATION

	OKA-MOTO	MORI	CHA	NAKAO	PROTOTYPE
1a	I	I	Pt	GI	HCV-1, HCV-H, HC-J1
1b	II	II	KI	GII	HCV-J, HCV-BK, HCV-T, HC-JK1, HC-J4, HCV-CHINA
1c					HC-G9
2a	III	III	K2a	GIII	HC-J6
2b	IV	IV	K2b	GIII	HC-J8

2c					S83, ARG6, ARG8, I10, T983
2d					NE92
3a	V	V	K3	GIV	BR36, BR56, HD10, N2L1, BR33, Ta, E-b1
3b		VI	K3	GIV	HCV-TR, Tb, NE137
3c					NE48
3d					NE274
3e					NE145
3f					NE125
4a					Z4, GB809-4
4b					Z1
4c					GB116, GB358, GB215, Z6, Z7
4d					DK13
4e					GB809-2, CAM600, CAM736
4f					CAM622, CAM627
4g					GB549
4h					GB438
4i					CAR4/1205
4j					CAR1/905
5a				GV	SA3, SA4, SA1, SA7, SA11, BE95
6a					HK1, HK2, HK3, HK4, VN11

Table 3 Overview of the known HCV types and subtypes classified according to the different authors.

The term "complement" refers to a nucleotide sequence which is complementary to an indicated sequence and which is able to hybridize to the indicated sequences.

The composition of the invention can comprise many combinations. By way of example, the composition of the invention can comprise:

- two (or more) nucleic acids from the same region or,
- two nucleic acids (or more), respectively from different regions, for the same isolate or for different isolates,
- or nucleic acids from the same regions and from at least two different regions

(for the same isolate or for different isolates).

The present invention relates particularly to a polynucleic acid as defined above having a sequence selected from any of SEQ ID NO 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39 , 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103 to 105, or a part of said polynucleic acid which is unique to any of the HCV subtypes or types as defined in Table 5, and which contains at least one nucleotide differing from known HCV polynucleic acids, or the complement thereof.

The present invention relates more particularly to a polynucleic acid as defined above, which codes for the 5' UR, the Core/E1, the NS4 or the NS5B region or a part thereof.

More particularly, the present invention relates to a polynucleic acid as defined above which is a cDNA sequence.

Also included within the present invention are sequence variants of the polynucleic acids as selected from any of the nucleotide sequences as given in any of the above given SEQ ID numbers with said sequence variants containing either deletion and/or insertions of one or more nucleotides, especially insertions or deletions of 1 or more codons, mainly at the extremities of oligonucleotides (either 3' or 5'), or substitutions of some non-essential nucleotides (i.e. nucleotides not essential to discriminate between different genotypes of HCV) by others (including modified nucleotides an/or inosine), for example, a type 1 or 2 sequence might be modified into a type 7 sequence by replacing some nucleotides of the type 1 or 2 sequence with type-specific nucleotides of type 7 as shown in for instance Figure 1 and 2.

Particularly preferred variant polynucleic acids of the present invention include also sequences which hybridise under stringent conditions with any of the polynucleic acid sequences of the present invention. Particularly, sequences which show a high degree of homology (similarity) to any of the polynucleic acids of the invention as described above. Particularly sequences which are at least 80%, 85%, 90%, 95% or more homologous to said polynucleic acid sequences of the invention. Preferably said sequences will have less than 20%, 15%, 10%, or 5% variation of the original nucleotides of said polynucleic acid sequence.

Polynucleic acid sequences according to the present invention which are

homologous to the sequences as represented by a SEQ ID NO can be characterized and isolated according to any of the techniques known in the art, such as amplification by means of sequence-specific primers, hybridization with sequence-specific probes under more or less stringent conditions, serological screening methods or via the LiPA typing system.

Other preferred variant polynucleic acids of the present invention include sequences which are redundant as a result of the degeneracy of the genetic code compared any of the above-given polynucleic acids of the present invention. These variant polynucleic acid sequences will thus encode the same amino acid sequence as the polynucleic acids they are derived from.

Also included within the scope of the present invention are 5' non-coding region sequences which can be readily obtained from type 1 subtype 1d, 1e, 1f or 1g isolates; type 2 subtype 2e, 2f, 2g, 2h, 2i, 2k or 2l isolates; type 3 subtype 3g isolates; type 4 subtype 4k, 4l or 4m isolates; type 7 subtype 7a, 7c or 7d isolates, type 9, type 10 or type 11 isolates discribed herein. Such sequences may contain type or subtype-specific motifs which can be employed for type and/or subtype-specific hybridization assays, e.g. such as described by Stuyver et al. (1993).

Polynucleic acid sequences of the genomes indicated above from regions not yet depicted in the present examples, figures and sequence listing can be obtained by any of the techniques known in the art, such as amplification techniques using suitable primers from the sequences of these new genomes given in Figure 1 of the present invention.

The present invention also relates to an oligonucleotide primer comprising part of a polynucleic acid as defined above, with said primer being able to act as a primer for specifically amplifying the nucleic acid of a certian HCV isolate belonging to the genotype from which the primer is derived.

The term "primer" refers to a single stranded DNA oligonucleotide sequence capable of acting as a point of initiation for synthesis of a primer extension product which is complementary to the nucleic acid strand to be copied. The length and the sequence of the primer must be such that they allow to prime the synthesis of the extension products. Preferably the primer is about 5-50 nucleotides. Specific length and sequence will depend on the complexity of the required DNA or RNA targets, as well as on the conditions of primer use such as temperature and ionic strength.

The fact that amplification primers do not have to match exactly with corresponding template sequence to warrant proper amplification is amply documented in the literature (Kwok et al., 1990).

The amplification method used can be either polymerase chain reaction (PCR; Saiki et al., 1988), ligase chain reaction (LCR; Landgren et al., 1988; Wu & Wallace, 1989; Barany, 1991), nucleic acid sequence-based amplification (NASBA; Guatelli et al., 1990; Compton, 1991), transcription-based amplification system (TAS; Kwoh et al., 1989), strand displacement amplification (SDA; Duck, 1990; Walker et al., 1992) or amplification by means of Qß replicase (Lizardi et al., 1988; Lomeli et al., 1989) or any other suitable method to amplify nucleic acid molecules using primer extension. During amplification, the amplified products can be conveniently labelled either using labelled primers or by incorporating labelled nucleotides. Labels may be isotopic (^{32}P, ^{35}S, etc.) or non-isotopic (biotin, digoxigenin, etc.). The amplification reaction is repeated between 20 and 70 times, advantageously between 25 and 45 times.

The present invention also relates to an oligonucleotide probe comprising part of a polynucleic acid as defined above, with said probe being able to act as a hybridization probe for specific detection and/or classification into types and/or subtypes of an HCV nucleic caid containing said nucleotide sequence, with said probe being possibly labelled or attached to a solid substrate.

The term "probe" refers to single stranded sequence-specific oligonucleotides which have a sequence which is complementary to the target sequence of the HCV genotype(s) to be detected.

Preferably, these probes are about 5 to 50 nucleotides long, more preferably from about 10 to 25 nucleotides.

The term "solid support" can refer to any substrate to which an oligonucleotide probe can be coupled, provided that it retains its hybridization characteristics and provided that the background level of hybridization remains low. Usually the solid substrate will be a microtiter plate, a membrane (e.g. nylon or nitrocellulose) or a microsphere (bead). Prior to application to the membrane or fixation it may be convenient to modify the nucleic acid probe in order to facilitate fixation or improve the hybridization efficiency. Such modifications may encompass homopolymer tailing, coupling with different reactive groups such as aliphatic

groups, NH_2 groups, SH groups, carboxylic groups, or coupling with biotin or haptens.

The present invention also relates to a diagnostic kit for use in determining the genotype of HCV, said kit comprising a primer as defined above.

The present invention also relates to a diagnostic kit for use in determining the genotype of HCV, said kit comprising a probe as defined above.

The present invention also relates to a diagnostic kit as defined above, wherein said probe(s) is(are) attached to a solid substrate.

The present invention also relates to a diagnostic kit as defined above, wherein a range of said probes is attached to specific locations on a solid substrate.

The present invention also relates to a diagnostic kit as defined above, wherein said solid support is a membrane strip and said probes are coupled to the membrane in the form of parallel lines.

The present invention also relates to a method for the detection of HCV nucleic acids present in a biological sample, comprising:

(i) possibly extracting sample nucleic acid,

(ii) amplifying the nucleic acid with at least one primer as defined above,

(iii) detecting the amplified nucleic acids.

The present invention also relates to a method for the detection of HCV nucleic acids present in a biological sample, comprising:

(i) possibly extracting sample nucleic acid,

(ii) possibly amplifying the nucleic acid with at least one primer as defiend above, or with a universal HCV primer,

(iii) hybridizing the nucleic acids of the biological sample, possibly under denatured conditions, at appropriate conditions with one or more probes as defined above, with said probes being preferably attached to a solid substrate,

(iv) possibly washing at appropriate conditions,

(v) detecting the hybrids formed.

The present invention also relates to a method for detecting the presence of one or more HCV genotypes present in a biological sample, comprising:

(i) possibly extracting sample nucleic acid,

(ii) specifically amplifying the nucleic acid with at least one primer as defined

above,

(iii) detecting said amplified nucleic acids.

The present invention also relates to a method for detecting the presence of one or more HCV genotypes present in a biological sample, comprising:

(i) possibly extracting sample nucleic acid,

(ii) possibly amplifying the nucleic acid with at least one primer as defined above or with a universal HCV primer,

(iii) hybridizing the nucleic acids of the biological sample, possibly under denatured conditions, at appropriate conditions with one or more probes as defined above, with said probes being preferably attached to a solid substrate,

(iv) possibly washing at appropriate conditions,

(v) detecting the hybrids formed,

(vi) inferring the presence of one or more HCV genotypes present from the observed hybridization pattern.

The present invention also relates to a method as defined above, wherein said probes are further characterized as defined above.

The present invention also relates to a method as defined above, wherein said nucleic acids are labelled during or after amplification.

Preferably, this technique could be performed in the 5' non-coding, Core or NS5B region.

The term "nucleic acid" can also be referred to as analyte strand and corresponds to a single- or double-stranded nucleic acid molecule. This analyte strand is preferentially positive- or negative stranded RNA, cDNA or amplified cDNA.

The term "biological sample" refers to any biological sample (tissue or fluid) containing HCV nucleic acid sequences and refers more particularly to blood serum or plasma samples.

The term "universal HCV primer" refers to oligonucleotide sequences complementary to any of the conserved regions of the HCV genome.

The expression "appropriate" hybridization and washing conditions are to be understood as stringent and are generally known in the art (e.g. Maniatis et al., Molecular Cloning: A Laboratory Manual, New York, Cold Spring Harbor Laboratory, 1982).

However, according to the hybridization solution (SSC, SSPE, etc.), these probes should be hybridized at their appropriate temperature in order to attain sufficient specificity.

The term "labelled" refers to the use of labelled nucleic acids. This may include the use of labelled nucleotides incorporated during the polymerase step of the amplification such as illustrated by Saiki et al. (1988) or Bej et al. (1990) or labelled primers, or by any other method known to the person skilled in the art.

The process of the invention comprises the steps of contacting any of the probes as defined above, with one of the following elements:

- either a biological sample in which the nucleic acids are made available for hybridization,
- or the purified nucleic acids contained in the biological sample
- or a single copy derived from the purified nucleic acids,
- or an amplified copy derived from the purified nucleic acids, with said elements or with said probes being attached to a solid substrate.

The expression "inferring the presence of one or more HCV genotypes present from the observed hybridization pattern" refers to the identification of the presence of HCV genomes in the sample by analyzing the pattern of binding of a panel of oligonucleotide probes. Single probes may provide useful information concerning the presence or absence of HCV genomes in a sample. On the other hand, the variation of the HCV genomes is dispersed in nature, so rarely is any one probe able to identify uniquely a specific HCV genome. Rather, the identity of an HCV genotype may be inferred from the pattern of binding of a panel of oligonucleotide probes, which are specific for (different) segments of the different HCV genomes. Depending on the choice of these oligonucleotide probes, each known HCV genotype will correspond to a specific hybridization pattern upon use of a specific combination of probes. Each HCV genotype will also be able to be discriminated from any other HCV genotype amplified with the same primers depending on the choice of the oligonucleotide probes. Comparison of the generated pattern of positively hybridizing probes for a sample containing one or more unkown HCV sequences to a scheme of expected hybridization patterns, allows one to clearly infer the HCV genotypes present in said sample.

The present invention thus relates to a method as defined above, wherein one

or more hybridization probes are selected from any of SEQ ID NO 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103 or 105 or sequence variants thereof as defined above.

In order to distinguish the amplified HCV genomes from each other, the target polynucleic acids are hybridized to a set of sequence-specific DNA probes targetting HCV genotypic regions (unique regions) located in the HCV polynucleic acids.

Most of these probes target the most type- or subtype-specific regions of HCV genotypes, but some can be caused to hybridize to more than one HCV genotype.

According to the hybridization solution (SSC, SSPE, etc.), these probes should be stringently hybridized at their appropriate temperature in order to attain sufficient specificity. However, by slightly modifying the DNA probes, either by adding or deleting one or a few nucleotides at their extremities (either 3' or 5'), or substituting some non-essential nucleotides (i.e. nucleotides not essential to discriminate between types) by others (including modified nucleotides or inosine) these probes or variants thereof can be caused to hybridize specifically at the same hybridization conditions (i.e. the same temperature and the same hybridization solution). Also changing the amount (concentration) of probe used may be beneficial to obtain more specific hybridization results. It should be noted in this context, that probes of the same length, regardless of their GC content, will hybridize specifically at approximately the same temperature in TMACl solutions (Jacobs et al., 1988).

Suitable assay methods for purposes of the present invention to detect hybrids formed between the oligonucleotide probes and the nucleic acid sequences in a sample may comprise any of the assay formats known in the art, such as the conventional dot-blot format, sandwich hybridization or reverse hybridization. For example, the detection can be accomplished using a dot blot format, the unlabelled amplified sample being bound to a membrane, the membrane being incorporated with at least one labelled probe under suitable hybridization and wash conditions, and the presence of bound probe being monitored.

An alternative and preferred method is a "reverse" dot-blot format, in which the amplified sequence contains a label. In this format, the unlabelled oligonucleotide probes are bound to a solid support and exposed to the labelled sample under appropriate stringent hybridization and subsequent washing conditions. It is to be

understood that also any other assay method which relies on the formation of a hybrid between the nucleic acids of the sample and the oligonucleotide probes according to the present invention may be used.

According to an advantageous embodiment, the process of detecting one or more HCV genotypes contained in a biological sample comprises the steps of contacting amplified HCV nucleic acid copies derived from the biological sample, with oligonucleotide probes which have been immobilized as parallel lines on a solid support.

According to this advantageous method, the probes are immobilized in a Line Probe Assay (LiPA) format. This is a reverse hybridization format (Saiki et al., 1989) using membrane strips onto which several oligonucleotide probes (including negative or positive control oligonucleotides) can be conveniently applied as parallel lines.

The invention thus also relates to a solid support, preferably a membrane strip, carrying on its surface, one or more probes as defined above, coupled to the support in the form of parallel lines.

The LiPA is a very rapid and user-friendly hybridization test. Results can be read after 4 hours. after the start of the amplification. After amplification during which usually a non-isotopic label is incorporated in the amplified product, and alkaline denaturation, the amplified product is contacted with the probes on the membrane and the hybridization is carried out for about 1 to 1,5 h hybridized polynucleic acid is detected. From the hybridization pattern generated, the HCV type can be deduced either visually, but preferably using dedicated software. The LiPA format is completely compatible with commercially available scanning devices, thus rendering automatic interpretation of the results very reliable. All those advantages make the LiPA format liable for the use of HCV detection in a routine setting. The LiPA format should be particularly advantageous for detecting the presence of different HCV genotypes.

The present invention also relates to a method for detecting and identifying novel HCV genotypes, different from the known HCV genomes, comprising the steps of:

- determining to which HCV genotype the nucleotides present in a biological sample belong, according to the process as defined above,
- in the case of observing a sample which does not generate a hybridization

pattern compatible with those defined in Table 3, sequencing the portion of the HCV genome sequence corresponding to the aberrantly hybridizing probe of the new HCV genotype to be determined.

The present invention also relates to a method for preparing a polynucleic acid according to the present invention. These methods include any method known in the art for preparing polynucleic acids (e.g. the phosphodiester method for synthesizing oligonucleotides as described by Agarwal et al. 1972, Agnew. Chem. Int. Ed. Engl. 11:451, the phosphotriester method of Hsiung et al. 1979, Nucleic Acid Res. 6:1371, or the automated diethylphosphoramidite method of Baeucage et al. 1981, Tetrahedron Letters 22:1859-1862.). Alternatively, the polynucleic acids of the present invention may be isolated fragments of naturally occuring or cloned DNA or RNA. In addition, the oligonucleotides according to the present invention may be synthesized automatically on commercial instruments sold by a variety of manufacturers.

The present invention particularly also relates to a polypeptide having an amino acid sequence encoded by a polynucleic acid as defined above, or a part thereof which is unique to at least one of the HCV subtypes or types as defined in Table 5, and which contains at least one amino acid differing from any of the known HCV types or subtypes, or an analog thereof being substantially homologous and biologically equivalent .

The term 'polypeptide' refers to a polymer of amino acids and does not refer to a specific length of the product; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide. This term also does not refer to or exclude post-expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like. Included within the definition are, for example, polypeptides containing one or more analogues of an amino acid (including, for example, unnatural amino acids, PNA, etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

The term "unique" is referred above.

By "biologically equivalent" as used throughout the specification and claims, it is meant that the compositions are immunogenically equivalent to the proteins (polypeptides) or peptides of the invention as defined above and below.

By "substantially homologous" as used throughout the ensuing specification and claims to describe proteins and peptides, it is meant a degree of homology in the amino acid sequence to the proteins or peptides of the invention. Preferably the degree of homology is in excess of 90, preferably in excess of 95, with a particularly preferred group of proteins being in excess of 99 homologous with the proteins or peptides of the invention.

The term "analog" as used throughout the specification or claims to describe the proteins or peptides of the present invention, includes any protein or peptide having an amino acid residue sequence substantially identical to a sequence specifically shown herein in which one or more residues have been conservatively substituted with a biologically equivalent residue. Examples of conservative substitutions include the substitution of one-polar (hydrophobic) residue such as isoleucine, valine, leucine or methionine for another, the substitution of one polar (hydrophillic) residue for another such as between arginine and lysine, between glutamine and asparagine, between glycine and serine, the substitution of one basic residue such as lysine, arginine or histidine for another, or the substitution of one acidic residue, such as aspartic acid or glutamic acid for another. Examples of allowable mutations acccording to the present inevntion can be found in Table 4.

The phrase "conservative substitution" also includes the use of a chemically derivatized residue in place of a non-derivatized residue provided that the resulting protein or peptide is biologically equivalent to theprotein or peptide of the invention.

"Chemical derivative" refers to a protein or peptide having one or more residues chemically derivatized by reaction of a functional side group. Examples of such derivatized molecules, include but are not limited to, those molecules in which free amino groups have been derivatized to form amine hydrochlorides, p-toluene sulfonyl groups, carbobenzoxy groups, t-butyloxycarbonyl groups, chloracetyl groups or formyl groups. Free carboxyl groups may be derivatized to form salts, methyl and ethyl esters or other types of esters or hydrazides. Free hydroxyl groups may be derivatized to form O-acyl or O-alkyl derivatives. The imidazole nitrogen of histidine may be derivatized to form N-imbenzylhistidine. Also included as chemical derivatives are those proteins or peptides which contain one or more naturally-occurring amino acid derivatives of the twenty standard amino acids. For examples : 4-hydroxyproline may be substituted for proline; 5-hydroxylysine may be substituted for lysine; 3-

methylhistidine may be substituted for histidine; homoserine may be substituted for serine; and ornithine may be substituted for lysine. The proteins or peptides of the present invention also include any protein or peptide having one or more additions and/or deletions or residues relative to the sequence of a peptide whose sequence is shown herein, so long as the peptide is biologically equivalent to the proteins or peptides of the invention.

It is to be noted that, at the level of the amino acid sequence, at least one amino acids difference (with respect to known HCV amino acid sequences) is sufficient to be part of the invention, which means that the polypeptides of the invention correspond to polynucleic acids having at least one nucleotide difference (with known HCV polynucleic acid sequences) involving an amino acid difference in the encoded polyprotein.

As the NS4 and the Core regions are known to contain several epitopes, for example characterized in patent application EP-A-0 489 968, and as the E1 protein is expected to be subject to immune attack as part of the viral envelope and expected to contain epitopes, the NS4, Core and E1 epitopes of the new types and subtypes disclosed herein will consistently differ from the epitopes present in previously known genotypes. This is examplified by the type-specificity of NS4 synthetic peptides as described in Simmonds et al. (1993c) and Stuyver et al. (1993b) and PCT/EP 94/01323 and the type-specificity of recombinant E1 proteins as described in Maertens et al. (1994).

The peptides according to the present invention contain preferably at least 3, preferably 4, 5 contiguous HCV amino acids, 6, 7 preferably however at least 8 contiguous HCV amino acids, at least 10 or at least 15 (for instance at least 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50 or more amino acids).

TABLE 4

Amino acids	Synonymous groups
Ser (S)	Ser, Thr, Gly, Asn
Arg (R)	Arg, His, Lys, Glu, Gln

Leu (L)	Leu; Ile, Met, Phe, Val, Tyr
Pro (P)	Pro, Ala, Thr, Gly
Thr (T)	Thr, Pro, Ser, Ala, Gly, His, Gln
Ala (A)	Ala, Pro, Gly, Thr
Val (V)	Val, Met, Ile, Tyr, Phe, Leu, Val
Gly (G)	Gly, Ala, Thr, Pro, Ser
Ile (I)	Ile, Met, Leu, Phe, Val, Ile, Tyr
Phe (F)	Phe, Met, Tyr, Ile, Leu, Trp, Val
Tyr (Y)	Tyr, Phe, Trp, Met, Ile, Val, Leu
Cys (C)	Cys, Ser, Thr, Met
His (H)	His, Gln, Arg, Lys, Glu, Thr
Gln (Q)	Gln, Glu, His, Lys, Asn, Thr, Arg
Asn (N)	Asn, Asp, Ser, Gln
Lys (K)	Lys, Arg, Glu, Gln, His
Asp (D)	Asp, Asn, Glu, Gln
Glu (E)	Glu, Gln, Asp, Lys, Asn, His, Arg
Met (M)	Met, Ile, Leu, Phe, Val

Table 4 Overview of the amino acid substitutions which could form the basis of analogs (muteins) as defined above

The polypeptides of the invention, and particularly the fragments, can be prepared by classical chemical synthesis.

The synthesis can be carried out in homogeneous solution or in solid phase.

For instance, the synthesis technique in homogeneous solution which can be used is the one described by Houbenweyl in the book entitled "Methode der organischen chemie" (Method of organic chemistry) edited by E. Wunsh, vol. 15-I et II. THIEME, Stuttgart 1974.

The polypeptides of the invention can also be prepared in solid phase according to the methods described by Atherton and Shepard in their book entitled "Solid phase peptide synthesis" (IRL Press, Oxford, 1989).

The polypeptides according to this invention can be prepared by means of recombinant DNA techniques as described by Maniatis et al., Molecular Cloning: A Laboratory Manual, New York, Cold Spring Harbor Laboratory, 1982).

The present invention relates particularly to a polypeptide as defined above, comprising in its amino acid sequence at least one of the following amino acid residues:

I15, C38, V44, A49, Q43, P49, Q55, A58, S60 or D60, E68 or V68, H70, A71 or Q71 or N71, D72, H81, H101, D106, S110, L130, I134, E135, L140, S148, T150 or E150, Q153, F155, D157, G160, E165, I169, F181, L186, T190, T192 or I192 or H192, I193, A195, S196, R197 or N197 or K197, Q199 or D199 or H199, N199, F200 or T200, A208, I213, M216 or S216, N217 or S217 or G217 or K217, T218, I219, A222, Y223, I230, W231 or L231, S232 or H232 or A232, Q233, E235 or L235, F236 or T236, F237, L240 or M240, A242, N244, N249, I250 or K250 or R250, A252 or C252, A254, I255 or V255, D256 or M256, E257, E260 or K260, R261, V268, S272 or R272, I285, G290 or F290, A291, A293 or L293 or W293, T294 or A294, S295, H295, K296 or E296, Y297 or M297, I299 or Y299, I300, S301, P316, S2646, A2648, G2649, A2650, V2652, Q2653, H2656 or L2656, D2657, F2659, K2663 or Q2663, A2667 or V2667, D2677, L2681, M2686 or Q2686 or E2686, A2692 or K2692, H2697, I2707, L2708 or Y2708, A2709, A2719 or M2719, F2727, T2728 or D2728, E2729, F2730 or Y2730, I2741, I2745, V2746 or E2746 or L2746 or K2746, A2748, S2749 or P2749, R2750, E2751, D2752 or N2752 or S2752 or T2752 or V2752 or I2752 or Q2752, S2753 or D2753 or G2753, D2754, A2755, L2756 or Q2756, or R2757,

with said notation being composed of a letter representing the amino acid residue by its one-letter code, and a number representing the amino acid numbering according to Kato et al., 1990 as shown in Table 1 (see also the numbering in Figures 2, 4 and 6),

or a part thereof which is unique to at least one of the HCV subtypes or types as defined in Table 5, and which contains at least one amino acid differing from any of the known HCV types or subtypes, or an analog thereof being substantially homologous and biologically equivalent to said polypeptide or part thereof.

These unique amino acid residues can be deduced from aligning the new HCV amino acid sequences as given in Figure 3 to all known HCV sequences. An alignment with the new sequences as represented in SEQ ID NO 1 to 106 is given in for instance Figures 2, 4 and 6. It should be clear that the alignments given in these figures may be completed with all known HCV sequences to illustrate that any of the above-given unique residues is indeed unique for at least one of the new HCV sequences of the present invention.

Within the group of unique and new amino acid residues of the present

invention, unique residues may be found which are specific for the following new types (subtypes) of HCV according to the HCV classification system used in the present invention: type 1 subtype 1d, 1e, 1f or 1g isolates; type 2 subtype 2e, 2f, 2g, 2h, 2i, 2k or 2l isolates; type 3 subtype 3g isolates; type 4 subtype 4k, 4l or 4m isolates; type 7 subtype 7a, 7c or 7d isolates, type 9, type 10 or type 11 isolates. In order to obtain these residues the alignments given in Figures 2, 4 and 6 may be used to deduce the type- and or subtype-specificity of any of the unique residues given above.

For example T190 (detected in subtype 1d) refers to a threomine at position 190 (see Figure 2). In other sequences only a serine (S190) or exceptionally an alanine (A190 in type 10a) can be detected.

The polypeptides according to this embodiment of the invention may be possibly labelled, or attached to a solid substrate, or coupled to a carrier molecule such as biotin, or mixed with a proper adjuvant all known in the art and according to the intended use (diagnostic, therapeutic or prophylactic).

The present invention also relates to a polypeptide as defined above, comprising in its amino acid sequence at least one of the sequences repesented by SEQ ID NO107 to 207 as listed above, or a part thereof which is unique to at least one of the HCV subtypes or types as defined in Table 5, or an analog thereof being substantially homologous and biologically equivalent to said polypeptide or part thereof.

The present invention relates also to a polypeptide having an amino acid sequence as represented in any of SEQ ID NO 1 to 106, or a part thereof which is unique to at least one of the HCV subtypes or types as defined in Table 5, or an analog thereof being substantially homologous and biologically equivalent to said polypeptide or part thereof.

The variable region in the core protein (V-CORE in Fig. 2) has been shown to be useful for serotyping (Machida et al., 1992). The sequence of the type 1 subtype 1d, 1e, 1f or 1g sequence; type 2 subtype 2e, 2f, 2g, 2h, 2i, 2k and 2l sequence; type 3 subtype 3g; type 4, subtype 4k, 4l or 4m sequence; type 7 (subtype 7a, 7c and 7d sequences), 9, 10 or 11 sequences of the present invention show type-specific features in this region. The peptide from amino acid 68 to 78 (V-core region) shows the following unique sequence for the sequences of the present invention (see

figure 2):

ARQSDGRSWAQ or ARRSEGRSWAQ as for subtype 1d (SEQ ID NO 107 and 108)

ERRPEGRSWAQ as for subtype 1e (SEQ ID NO 109)

ARRPEGRSWAQ as for subtype 1f (SEQ ID NO 110)

DRRTTGKSWGR as for subtype 2k (SEQ ID NO 111)

DRRATGRSWGR as for subtype 2e (SEQ ID NO 112)

DRRATGKSWGR as for subtype 2f (SEQ ID NO 113)

VRQPTGRSWGQ as for type 9 (SEQ ID NO 114)

VRHQTGRTWAQ as for subtype 7a and 7c (SEQ ID NO 115)

VRQNQGRTWAQ as for subtype 7d (SEQ ID NO 116)

ARRTEGRSWAQ as for type 10 (SEQ ID NO 117)

VRRTTGRXXXX or VRRTTGRTWAQ as for type 11 (SEQ ID NO 118 and 119)

Five type-specific variable regions (V1 to V5) can be identified after aligning E1 amino acid sequences of the genotypes of the present invention to the genotypes already known, as shown in Figure 2.

Region V1 encompasses amino acids 192 to 203, this is the amino-terminal 10 amino acids of the E1 protein. The following unique sequences as shown in Fig. 2 can be deduced:

HEVRNASGVYHV or HEVRNASGVYHL as for subtype 1d, (SEQ ID NO 120 and 121)

YEVHSTTDGYHV as for subtype 1f (SEQ ID NO 122)

VEVKNTSQAYMA as for subtype 2e (SEQ ID NO 123)

IQVKNNSHFYMA as for subtype 2f (SEQ ID NO 124)

VQVKNTSTMYMA as for subtype 2g (SEQ ID NO 125)

VQVKNTSHSYMV as for subtype 2h (SEQ ID NO 126)

VQVANRSGSYMV as for subtype 2i (SEQ ID NO 127)

VEIKNTXNTYVL or VEIKNTSNTYVL as for subtype 2k (SEQ ID NO 128 and 129)

INYRNVSGIYYV or INYRNTSGIYHV or INYHNTSGIYHI or TNYRNVSGIYHV a for subtype 4k (SEQ ID NO 130, 131, 132 or 133)

QHYRNVSGIYHV as for subtype 4l (SEQ ID NO 134)

IQVKNASGIYHL as for type 9 (SEQ ID NO 135)

AHYTNKSGLYHL as for subtype 7c (SEQ ID NO 136)

LNYANKSGLYHL as for subtype 7d (SEQ ID NO 137)

LEYRNASGLYMV as for type 10 (SEQ ID NO 138)

Region V2 encompasses amino acids 213 to 223. The following unique sequences can be found in the V2 region as shown in Figure 2:

IYEMDGMIMHY or IYEMSGMILHA as for subtype 1d, (SEQ ID NO 139 and 140)

VYEAKDIILHT as for subtype 1f (SEQ ID NO 141)

VWQLXDAVLHV as for subtype 2e (SEQ ID NO 142)

VWQLRDAVLHV as for subtype 2f (SEQ ID NO 143)

IWQMQGAVLHV as for subtupe 2g (SEQ ID NO 144)

VWQLKDAVLHV as for subtype 2h (SEQ ID NO 145)

VWQLEEAVLHV as for subtype 2i (SEQ ID NO 146)

TWQLXXAVLHV as for subtype 2k (SEQ ID NO 147)

VYEADHHILHL or VYEADHHILAL or VFEADHHILHL as for subtype 4k (SEQ ID NO 148, 149 and 150)

VYESDHHILHL as for subtype 4l (SEQ ID NO 151)

VFEAETMILHL as for type 9 (SEQ ID NO 152)

VYEAETLILHL as for subtype 7c (SEQ ID NO 153)

VYEANGMILHL as for subtype 7d (SEQ ID NO 154)

VYEAGDIILHL as for type 10. (SEQ ID NO 155)

Region V3 encompasses the amino acids 230 to 242. The following unique V3 region sequences can be deduced from Figure 2:

VREDNHLRCWMAL or VRENNSSRCWMAL as for subtype 1d (SEQ ID NO 156 and 157)

IREGNISRCWVLP as for subtype 1f (SEQ ID NO 158)

ENSSGRFHCWIPI as for subtype 2e (SEQ ID NO 159)

ERSGNRTFCWTAV as for subtype 2f (SEQ ID NO 160)

ELQGNKSRCWIPV as for subtype 2g (SEQ ID NO 162)

ERHQNQSRCWIPV as for subtype 2h (SEQ ID NO 163)

EWKDNTSRCWIPV as for subtype 2i (SEQ ID NO 164)

EREGNSSRCWIPV as for subtype 2k (SEQ ID NO 165)

VREGNQSRCWVAL or VRTGNQSRCWVAL or VRVGNQSSCWVAL or VRVGNQSRCWVAL or VKEGNHSRCWVAL as for subtype 4k (SEQ ID NO 166, 167, 168 or 169)

VKTGNTSRCWVAL as for subtype 4l (SEQ ID NO 170)

IKAGNESRCWLPV as for type 9 (SEQ ID NO 171)

VKXXNQSRCWVQA as for subtype 7c (SEQ ID NO 172)

VKTGNLTKCWLSA as for subtype 7d (SEQ ID NO 173)

VRSGNTSRCWIPV as for type 10 (SEQ ID NO 174)

Region V4 encompasses the amino acids 248 to 257. The following unique V4 region sequences can be deduced from figure 2:

VKNASVPTAA or VKDANVPTAA as for subtype 1d (SEQ ID NO 175 and 176)

ARIANAPIDE as for subtype 1f (SEQ ID NO 177)

VSKPGALTKG as for subtype 2e (SEQ ID NO 178)

VSRPGALTRG as for subtype 2f (SEQ ID NO 179)

VNQPGALTRG as for subtype 2g (SEQ ID NO 180)

VSQPGALTRG as for subtype 2h (SEQ ID NO 181)

VSQPGALTKG as for subtype 2i (SEQ ID NO 182)

VSRPGALTEG as for subtype 2k (SEQ ID NO 183)

APYIGAPLES or APYTAAPLES as for subtype 4k (SEQ ID NO 184 and 185)

APILSAPLMS as for subtype 4l (SEQ ID NO 186)

VPNSSVPIHG as for type 9 (SEQ ID NO 187)

VPNASTPVTG as for subtype 7c (SEQ ID NO 188)

VQNASVSIRG as for subtype 7d (SEQ ID NO 189)

VKSPCAATAS as for type 10 (SEQ ID NO 190)

Region V5 encompasses the amino acids 294 to 303. The following unique V5 region peptides can be deduced from figure 2:

SPRMHHTTQE or SPRLYHTTQE as for subtype 1d (SEQ ID NO 191 and 192)

TSRRHWTVQD as for subtype 1f (SEQ ID NO 193)

APKRHYFVQE as for subtype 2e (SEQ ID NO 194)

SPQYHTFVQE as for subtype 2f (SEQ ID NO 195)

SPQHHNFSQD as for subtype 2g (SEQ ID NO 196)

SPQHHIFVQD as for subtype 2h (SEQ ID NO 197)

SPEHHHFVQD as for subtype 2k (SEQ ID NO 198)

RPRRHWTTQD or RPRRHWTAQD or QPRRHWTTQD or RPRRHWTTQE as for subtype 4k (SEQ ID NO 199, 200, 201 or 202)

QPRRHWTVQD as for subtype 4l (SEQ ID NO 203)

RPKYHQVTQD as for type 9 (SEQ ID NO 204)

RPRMHQVVQE as for subtype 7c (SEQ ID NO 205)

RPRMYEIAQD as for subtype 7d (SEQ ID NO 206)

RHRQHWTVQD as for type 10 (SEQ ID NO 207)

The above given list of peptides are particularly useful for treatment and vaccine and diagnostic development.

Also comprised in the present invention is any synthetic peptide (see below) or polypeptide containing at least an epitope derived from the above-defined peptides in their peptidic chain. Also comprised within the present invention is any synthetic peptide or polypeptide comprising at least 6, 7, 8, or 9 contiguous amino acids derived from the above-defined peptides in their peptidic chain.

As used herein, 'epitope' or 'antigenic determinant' means an amino acid sequence that is immunoreactive. Generally an epitope consists of at least 3 to 4 amino acids, and more usually, consists of at least 5 or 6 amino acids, sometimes the epitope consists of about 7 to 8, or even about 10 amino acids.

The present invention particularly relates to any peptide (see below) or polypeptide contained in any of the amino acid sequences as represented in SEQ ID NO 2, 4, 7, 9, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104 or 106 (see Table 5 and Figure 3, Examples section).

The present invention also relates to a recombinant polypeptide encoded by a polynucleic acid as defined above, or a part thereof which is unique to any of the HCV subtypes or types as defined in Table 5, or an analog thereof being substantially homologous and biologically equivalent to said polypeptide.

The present invention also relates to a recombinant expression vector comprising a polynucleic acid or a part thereof as defined above, operably linked to prokaryotic, eukaryotic or viral transcription and translation control elements.

In general said recombinant vector will comprise a vector sequence, an appropriate prokaryotic, eukaryotic or viral promoter sequence followed by the nucleotide sequences as defined above, with said recombinant vector allowing the expression of any one of the polypeptides as defined above in a prokaryotic, or eukaryotic host or in living mammals when injected as naked DNA, and more particularly a recombinant vector allowing the expression of any of the new HCV sequences of the invention spanning particularly the following amino acid positions:

- a polypeptide starting in the region between positions 1 and 10 and ending at any position in the region between positions 70 and 420, more particularly a polypeptide spanning positions 1 to 70, 1 to 85, positions 1 to 120, positions 1 to 150, positions 1 to 191, or positions 1 to 200, for expression of the Core protein, and a polypeptide spanning positions 1 to 263, positions 1 to 326, positions 1 to 383, or positions 1 to 420 for expression of the Core and E1 protein;
- a polypeptide starting at any position in the region between positions 117 and 192, and ending at any position in the region between positions 263 and 420, for expression of E1, or forms that have the hydrophobic region deleted (positions 264 to 293 plus or minus 8 amino acids);
- a polypeptide starting at any position in the region between positions 1556 and 1688, and ending at any position in the region between positions 1739 and 1764, for expression of NS4, more particularly ;a polypeptide starting at position 1658 and ending at position 1711, for expression of NS4a antigen, and more particularly, a polypeptide starting at position 1712 and ending in the region between positions 1743 and 1972 (for instance 1712-1743, 1712-1764, 1712-1782, 1712-1972, 1712-1782, 1712-1902), for expression of NS4b antigen or parts thereof.

Any other HCV vector construction known in the art may also be used for the recombinant polypeptides of the present invention.

Also any of the known purification methods for recombinant proteins may be used for the production of the recombinant polypeptides of the present invention, particularly the HCV recombinant polypeptide purification methods as disclosed in PCT/EP 95/03031 in name of Innogenetics N.V.

The term "vector" may comprise a plasmid, a cosmid, a phage, or a virus or

a transgenic animal. Particularly useful for vaccine development may be BCG or adenoviral vectors, as well as avipox recombinant viruses.

The present invention also relates to a method for the production of a recombinant polypeptide as defined above, comprising:

- transformation of an appropriate cellular host with a recombinant vector, in which a polynucleic acid or a part thereof according to as defined above has been inserted under the control of appropriate regulatory elements,
- culturing said transformed cellular host under conditions enabling the expression of said insert, and,
- harvesting said polypeptide.

The term 'recombinantly expressed' used within the context of the present invention refers to the fact that the proteins of the present invention are produced by recombinant expression methods be it in prokaryotes, or lower or higher eukaryotes as discussed in detail below.

The term 'lower eukaryote' refers to host cells such as yeast, fungi and the like. Lower eukaryotes are generally (but not necessarily) unicellular. Preferred lower eukaryotes are yeasts, particularly species within Saccharomyces, Schizosaccharomyces, Kluveromyces, Pichia (e.g. Pichia pastoris), Hansenula (e.g. Hansenula polymorpha), Yarowia, Schwaniomyces, Schizosaccharomyces, Zygosaccharomyces and the like. Saccharomyces cerevisiae, S. carlsbergensis and K. lactis are the most commonly used yeast hosts, and are convenient fungal hosts.

The term 'prokaryotes' refers to hosts such as E.coli, Lactobacillus, Lactococcus, Salmonella, Streptococcus, Bacillus subtilis or Streptomyces. Also these hosts are contemplated within the present invention.

The term 'higher eukaryote' refers to host cells derived from higher animals, such as mammals, reptiles, insects, and the like. Presently preferred higher eukaryote host cells are derived from Chinese hamster (e.g. CHO), monkey (e.g. COS and Vero cells), baby hamster kidney (BHK), pig kidney (PK15), rabbit kidney 13 cells (RK13), the human osteosarcoma cell line 143 B, the human cell line HeLa and human hepatoma cell lines like Hep G2, and insect cell lines (e.g. Spodoptera frugiperda). The host cells may be provided in suspension or flask cultures, tissue cultures, organ cultures and the like. Alternatively the host cells may also be transgenic animals.

The term 'recombinant polynucleotide or nucleic acid' intends a polynucleotide

or nucleic acid of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation : (1) is not associated with all or a portion of a polynucleotide with which it is associated in nature, (2) is linked to a polynucleotide other than that to which it is linked in nature, or (3) does not occur in nature.

The term 'recombinant host cells', 'host cells', 'cells', 'cell lines', 'cell cultures', and other such terms denoting microorganisms or higher eukaryotic cell lines cultured as unicellular entities refer to cells which can be or have been, used as recipients for a recombinant vector or other transfer polynucleotide, and include the progeny of the original cell which has been transfected. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement as the original parent, due to natural, accidental, or deliberate mutation.

The term 'replicon' is any genetic element, e.g., a plasmid, a chromosome, a virus, a cosmid, etc., that behaves as an autonomous unit of polynucleotide replication within a cell; i.e., capable of replication under its own control.

The term 'vector' is a replicon further comprising sequences providing replication and/or expression of a desired open reading frame.

The term 'control sequence' refers to polynucleotide sequences which are necessary to effect the expression of coding sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include promoter, ribosomal binding site, splicing sites and terminators; in eukaryotes, generally, such control sequences include promoters, splicing sites, terminators and, in some instances, enhancers. The term 'control sequences' is intended to include, at a minimum, all components whose presence is necessary for expression, and may also include additional components whose presence is advantageous, for example, leader sequences which govern secretion.

The term 'promoter' is a nucleotide sequence which is comprised of consensus sequences which allow the binding of RNA polymerase to the DNA template in a manner such that mRNA production initiates at the normal transcription initiation site for the adjacent structural gene.

The expression 'operably linked' refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their

intended manner. A control sequence 'operably linked' to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences.

The segment of the HCV cDNA encoding the desired sequence inserted into the vector sequence may be attached to a signal sequence. Said signal sequence may be that from a non-HCV source, e.g. the IgG or tissue plasminogen activator (tpa) leader sequence for expression in mammalian cells, or the α-mating factor sequence for expression into yeast cells, but particularly preferred constructs according to the present invention contain signal sequences appearing in the HCV genome before the respective start points of the proteins.

A variety of vectors may be used to obtain recombinant expression of HCV single or specific oligomeric envelope proteins of the present invention. Lower eukaryotes such as yeasts and glycosylation mutant strains are typically transformed with plasmids, or are transformed with a recombinant virus. The vectors may replicate within the host independently, or may integrate into the host cell genome.

Higher eukaryotes may be transformed with vectors, or may be infected with a recombinant virus, for example a recombinant vaccinia virus. Techniques and vectors for the insertion of foreign DNA into vaccinia virus are well known in the art, and utilize, for example homologous recombination. A wide variety of viral promoter sequences, possibly terminator sequences and poly(A)-addition sequences, possibly enhancer sequences and possibly amplification sequences, all required for the mammalian expression, are available in the art. Vaccinia is particularly preferred since vaccinia halts the expression of host cell proteins. Vaccinia is also very much preferred since it allows the expression of f.i. E1 and E2 proteins of HCV in cells or individuals which are immunized with the live recombinant vaccinia virus. For vaccination of humans the avipox and Ankara Modified Virus (AMV) are particularly useful vectors.

Also known are insect expression transfer vectors derived from baculovirus Autographa californica nuclear polyhedrosis virus (AcNPV), which is a helper-independent viral expression vector. Expression vectors derived from this system usually use the strong viral polyhedrin gene promoter to drive the expression of heterologous genes. Different vectors as well as methods for the introduction of heterologous DNA into the desired site of baculovirus are available to the man skilled

in the art for baculovirus expression. Also different signals for posttranslational modification recognized by insect cells are known in the art.

The present invention also relates to a host cell transformed with a recombinant vector as defined above.

The present invention also relates to a method for detecting antibodies to HCV present in a biological sample, comprising:

(i) contacting the biological sample to be analysed for the presence of HCV with a polypeptide as defined above,

(ii) detecting the immunological complex formed between said antibodies and said polypeptide.

The present invention also relates to a method for HCV typing, comprising:

(i) contacting the biological sample to be analysed for the presence of HCV with a polypeptide as defined above,

(ii) detecting the immunological complex formed between said antibodies and said polypeptide.

The present invention also relates to a diagnostic kit for use in detecting the presence of HCV, said kit comprising at least one polypeptide as defined above, with said polypeptide being preferably bound to a solid support.

The present invention also relates to a diagnostic kit for HCV typing, said kit comprising at least one polypeptide as defined above, with said polypeptide being preferably bound to a solid support.

The present invention also relates to diagnostic kit according as defined above, said kit comprising a range of said polypeptides which are attached to specific locations on a solid substrate.

The present invention also relates to a diagnostic kit as defined above, wherein said solid support is a membrane strip and said polypeptides are coupled to the membrane in the form of parallel lines.

The immunoassay methods according to the present invention may utilize antigens from the different domains of the new and unique polypeptide sequences of the present invention that maintain linear (in case of peptides) and conformational epitopes (in case of polypeptides) recognized by antibodies in the sera from individuals infected with HCV. It is within the scope of the invention to use for instance single or specific oligomeric antigens, dimeric antigens, as well as

combinations of single or specific oligomeric antigens. The HCVantigens of the present invention may be employed in virtually any assay format that employs a known antigen to detect antibodies. Of course, a format that denatures the HCV conformational epitope should be avoided or adapted. A common feature of all of these assays is that the antigen is contacted with the body component suspected of containing HCV antibodies under conditions that permit the antigen to bind to any such antibody present in the component. Such conditions will typically be physiologic temperature, pH and ionic strenght using an excess of antigen. The incubation of the antigen with the specimen is followed by detection of immune complexes comprised of the antigen.

Design of the immunoassays is subject to a great deal of variation, and many formats are known in the art. Protocols may, for example, use solid supports, or immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide; the labels may be, for example, enzymatic, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays which amplify the signals from the immune complex are also known; examples of which are assays which utilize biotin and avidin or streptavidin, and enzyme-labeled and mediated immunoassays, such as ELISA assays.

The immunoassay may be, without limitation, in a heterogeneous or in a homogeneous format, and of a standard or competitive type. In a heterogeneous format, the polypeptide is typically bound to a solid matrix or support to facilitate separation of the sample from the polypeptide after incubation. Examples of solid supports that can be used are nitrocellulose (e.g., in membrane or microtiter well form), polyvinyl chloride (e.g., in sheets or microtiter wells), polystyrene latex (e.g., in beads or microtiter plates, polyvinylidine fluoride (known as Immunolon™), diazotized paper, nylon membranes, activated beads, and Protein A beads. For example, Dynatech Immunolon™ 1 or Immunlon™ 2 microtiter plates or 0.25 inch polystyrene beads (Precision Plastic Ball) can be used in the heterogeneous format. The solid support containing the antigenic polypeptides is typically washed after separating it from the test sample, and prior to detection of bound antibodies. Both standard and competitive formats are know in the art.

In a homogeneous format, the test sample is incubated with the combination of antigens in solution. For example, it may be under conditions that will precipitate

any antigen-antibody complexes which are formed. Both standard and competitive formats for these assays are known in the art.

In a standard format, the amount of HCV antibodies in the antibody-antigen complexes is directly monitored. This may be accomplished by determining whether labeled anti-xenogeneic (e.g. anti-human) antibodies which recognize an epitope on anti-HCV antibodies will bind due to complex formation. In a competitive format, the amount of HCV antibodies in the sample is deduced by monitoring the competitive effect on the binding of a known amount of labeled antibody (or other competing ligand) in the complex.

Complexes formed comprising anti-HCV antibody (or in the case of competitive assays, the amount of competing antibody) are detected by any of a number of known techniques, depending on the format. For example, unlabeled HCV antibodies in the complex may be detected using a conjugate of anti-xenogeneic Ig complexed with a label (e.g. an enzyme label).

In an immunoprecipitation or agglutination assay format the reaction between the HCV antigens and the antibody forms a network that precipitates from the solution or suspension and forms a visible layer or film of precipitate. If no anti-HCV antibody is present in the test specimen, no visible precipitate is formed.

There currently exist three specific types of particle agglutination (PA) assays. These assays are used for the detection of antibodies to various antigens when coated to a support. One type of this assay is the hemagglutination assay using red blood cells (RBCs) that are sensitized by passively adsorbing antigen (or antibody) to the RBC. The addition of specific antigen antibodies present in the body component, if any, causes the RBCs coated with the purified antigen to agglutinate.

To eliminate potential non-specific reactions in the hemagglutination assay, two artificial carriers may be used instead of RBC in the PA. The most common of these are latex particles. However, gelatin particles may also be used. The assays utilizing either of these carriers are based on passive agglutination of the particles coated with purified antigens.

The HCV antigens of the present invention comprised of conformational epitopes will typically be packaged in the form of a kit for use in these immunoassays. The kit will normally contain in separate containers the native HCV antigen, control antibody formulations (positive and/or negative), labeled antibody

when the assay format requires the same and signal generating reagents (e.g. enzyme substrate) if the label does not generate a signal directly. The native HCV antigen may be already bound to a solid matrix or separate with reagents for binding it to the matrix. Instructions (e.g. written, tape, CD-ROM, etc.) for carrying out the assay usually will be included in the kit.

Immunoassays that utilize the native HCV antigen are useful in screening blood for the preparation of a supply from which potentially infective HCV is lacking. The method for the preparation of the blood supply comprises the following steps. Reacting a body component, preferably blood or a blood component, from the individual donating blood with HCV polypeptides of the present invention to allow an immunological reaction between HCV antibodies, if any, and the HCV antigen. Detecting whether anti-HCV antibody - HCV antigen complexes are formed as a result of the reacting. Blood contributed to the blood supply is from donors that do not exhibit antibodies to the native HCV antigens.

In cases of a positive reactivity to the HCV antigen, it is preferable to repeat the immunoassay to lessen the possibility of false positives. For example, in the large scale screening of blood for the production of blood products (e.g. blood transfusion, plasma, Factor VIII, immunoglobulin, etc.) 'screening' tests are typically formatted to increase sensitivity (to insure no contaminated blood passes) at the expense of specificity; i.e. the false-positive rate is increased. Thus, it is typical to only defer for further testing those donors who are 'repeatedly reactive'; i.e. positive in two or more runs of the immunoassay on the donated sample. However, for confirmation of HCV-positivity, the 'confirmation' tests are typically formatted to increase specificity (to insure that no false-positive samples are confirmed) at the expense of sensitivity.

The solid phase selected can include polymeric or glass beads, nitrocellulose, microparticles, microwells of a reaction tray, test tubes and magnetic beads. The signal generating compound can include an enzyme, a luminescent compound, a chromogen, a radioactive element and a chemiluminescent compound. Examples of enzymes include alkaline phosphatase, horseradish peroxidase and beta-galactosidase. Examples of enhancer compounds include biotin, anti-biotin and avidin. Examples of enhancer compounds binding members include biotin, anti-biotin and avidin. In order to block the effects of rheumatoid factor-like substances, the

test sample is subjected to conditions sufficient to block the effect of rheumatoid factor-like substances. These conditions comprise contacting the test sample with a quantity of anti-human IgG to form a mixture, and incubating the mixture for a time and under conditions sufficient to form a reaction mixture product substantially free of rheumatoid factor-like substance.

The present invention particularly relates to an immunoassay format in which the polypeptides (or peptides) of the invention are coupled to a membrane in the form of parallel lines . This assay format is particularly advantageous for HCV typing purposes.

The present invention also relates to a pharmaceutical composition comprising at least one (recombinant) polypeptides as defined above and a suitable excipient, diluent or carrier.

The present invention also relates to a method of preventing HCV infection, comprising administering the pharmaceutical composition as defined above to a mammal in effective amount to stimulate the production of protective antibody or protective T-cell response.

The present invention relates to the use of a composition as defined above in a method for preventing HCV infection.

The present invention further relates to a vaccine for immunizing a mammal against HCV infection, comprising at least one (recombinant) polypeptide as defined above, in a pharmaceutically acceptable carrier.

The term 'immunogenic' refers to the ability of a substance to cause a humoral and/or cellular response, whether alone or when linked to a carrier, in the presence or absence of an adjuvant. 'Neutralization' refers to an immune response that blocks the infectivity, either partially or fully, of an infectious agent. A 'vaccine' is an immunogenic composition capable of eliciting protection against HCV, whether partial or complete. A vaccine may also be useful for treatment of an individual, in which case it is called a therapeutic vaccine.

The term 'therapeutic' refers to a composition capable of treating HCV infection. The term 'effective amount' refers to an amount of epitope-bearing polypeptide sufficient to induce an immunogenic response in the individual to which it is administered, or to otherwise detectably immunoreact in its intended system (e.g., immunoassay). Preferably, the effective amount is sufficient to effect

treatment, as defined above. The exact amount necessary will vary according to the application. For vaccine applications or for the generation of polyclonal antiserum / antibodies, for example, the effective amount may vary depending on the species, age, and general condition of the individual, the severity of the condition being treated, the particular polypeptide selected and its mode of administration, etc. It is also believed that effective amounts will be found within a relatively large, non-critical range. An appropriate effective amount can be readily determined using only routine experimentation. Preferred ranges of proteins for prophylaxis of HCV disease are 0.01 to 100 μg/dose, preferably 0.1 to 50 μg/dose. Several doses may be needed per individual in order to achieve a sufficient immune response and subsequent protection against HCV disease.

The present invention also relates to a vaccine as defined above, comprising at least one (recombinant) polypeptide as defined above, with said polypeptide being unique for at least one of the subtypes or types as defined above.

Said vaccine compositions may include prophylactic as well as therapeutic vaccine compositions.

Pharmaceutically acceptable carriers include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers; and inactive virus particles. Such carriers are well known to those of ordinary skill in the art.

Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to : aluminim hydroxide (alum), N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP) as found in U.S. Patent No. 4,606,918, N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE) and RIBI, which contains three components extracted from bacteria, monophosphoryl lipid A, trehalose dimycolate, and cell wall skeleton (MPL + TDM + CWS) in a 2% squalene/Tween 80 emulsion. Any of the 3 components MPL, TDM or CWS may also be used alone or combined 2 by 2. Additionally, adjuvants such as Stimulon (Cambridge Bioscience, Worcester, MA)

Immunogenic compositions used as vaccines comprise a 'sufficient amount' or 'an immunologically effective amount' of the proteins of the present invention, as well as any other of the above mentioned components, as needed. 'Immunologically effective amount', means that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment, as defined above. This amount varies depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated (e.g. nonhuman primate, primate, etc.), the capacity of the individual's immune system to synthesize antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, the strain of infecting HCV, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials. Usually, the amount will vary from 0.01 to 1000 μg/dose, more particularly from 0.1 to 100 μg/dose.

The proteins of the invention may also serve as vaccine carriers to present homologous (e.g. T cell epitopes or B cell epitopes fromfor istance the core,E1, E2, NS2, NS3, NS4 or NS5 regions) or heterologous (non-HCV) haptens, in the same manner as Hepatitis B surface antigen (see European Patent Application 174,444). In this use, envelope proteins provide an immunogenic carrier capable of stimulating an immune response to haptens or antigens conjugated to the aggregate. The antigen may be conjugated either by conventional chemical methods, or may be cloned into the gene encoding E1 and/or E2 at a location corresponding to a hydrophilic region

of the protein. Such hydrophylic regions include the V1 region (encompassing amino acid positions 191 to 202), the V2 region (encompassing amino acid positions 213 to 223), the V3 region (encompassing amino acid positions 230 to 242), the V4 region (encompassing amino acid positions 230 to 242), the V5 region (encompassing amino acid positions 294 to 303) and the V6 region (encompassing amino acid positions 329 to 336). Another useful location for insertion of haptens is the hydrophobic region (encompassing approximately amino acid positions 264 to 293). It is shown in the present invention that this region can be deleted without affecting the reactivity of the deleted E1 protein with antisera. Therefore, haptens may be inserted at the site of the deletion.

The immunogenic compositions are conventionally administered parenterally, typically by injection, for example, subcutaneously or intramuscularly. Additional formulations suitable for other methods of administration include oral formulations and suppositories. Dosage treatment may be a single dose schedule or a multiple dose schedule. The vaccine may be administered in conjunction with other immunoregulatory agents.

The administration of the immunogen(s) of the present invention may be for either a prophylactic or therapeutic purpose. When provided prophylactically, the immunogen(s) is provided in advance of any exposure to HCV or in advance of any symptom of any symptoms due to HCV infection. The prophylactic administration of the immunogen serves to prevent or attenuate any subsequent infection of HCV in a mammal. When provided therapeutically, the immunogen(s) is provided at (or shortly after) the onset of the infection or at the onset of any symptom of infection or disease caused by HCV. The therapeutic administration of the immunogen(s) serves to attenuate the infection or disease.

In addition to use as a vaccine, the compositions can be used to prepare antibodies to HCV (E1) proteins. The antibodies can be used directly as antiviral agents. To prepare antibodies, a host animal is immunized using the E1 proteins native to the virus particle bound to a carrier as described above for vaccines. The host serum or plasma is collected following an appropriate time interval to provide a composition comprising antibodies reactive with the (E1) protein of the virus particle. The gamma globulin fraction or the IgG antibodies can be obtained, for example, by use of saturated ammonium sulfate or DEAE Sephadex, or other

techniques known to those skilled in the art. The antibodies are substantially free of many of the adverse side effects which may be associated with other anti-viral agents such as drugs.

The present invention also relates particularly to a peptide corresponding to an amino acid sequence encoded by at least one of the HCV genomic sequences as defined above, with said peptide being unique to any of the HCV subtypes or types as defined in Table 5, and which contains at least one amino acid differing from any of the known HCV types or subtypes, or an analog thereof being substantially homologous and biologically equivalent.

The present invention relates particularly to a peptide comprising at least one unique epitope of the new sequences of the invention as represented in SEQ ID NO 1 to 106.

The present invention relates also particularly to a peptide comprising in its sequence a unique amino acid residue of the invention as defined above.

The present invention relates particularly to a peptide which is biotinylated as explained in WO 93/18054.

All the embodiments (immunoassay formats, vaccines, compositions, uses, etc.) illustrated for the polypeptides of the invention as above also relate to the peptides of the invention.

The present invention also relates to a method for detecting antibodies to HCV present in a biological sample, comprising:

(i) contacting the biological sample to be analysed for the presence of HCV with a peptide as defined above,

(ii) detecting the immunological ccomplex formed between said antibodies and said peptide.

The present invention also relates to a method for HCV typing, comprising:

(i) contacting the biological sample to be analysed for the presence of HCV with a peptide as defined above,

(ii) detecting the immunological ccomplex formed between said antibodies and said peptide.

The present invention also relates to a diagnostic kit for use in detecting the presence of HCV, said kit comprising at least one peptide as defined above, with said peptide being preferably bound to a solid support.

The present invention also relates to a diagnostic kit for HCV typing, said kit comprising at least one peptide as defined above, with said peptide being preferably bound to a solid support.

The present invention also relates to a diagnostic kit as defined above, wherein said peptides are selected from the following:

- at least one NS4 peptide,
- at least one NS4 peptide and at least one Core peptide,
- at least one NS4 peptide and at least one Core peptide and at least one E1 peptide,
- at least one NS4 peptide and at least one E1 peptide.

The present invention also relates to a diagnostic kit as defined above, said kit comprising a range of said peptides which are attached to specific locations on a solid substrate.

The present invention also relates to a diagnostic kit as defined above, wherein said solid support is a membrane strip and said peptides are coupled to the membrane in the form of parallel lines.

The present invention also relates to a pharmaceutical composition comprising at least one as defined above and a suitable excipient, diluent or carrier.

the present invention also relates to a method of preventing HCV infection, comprising administering the pharmaceutical composition as defined above to a mammal in effective amount to stimulate the production of protective antibody or protective T-cell response.

The present invention also relates to the use of a composition as defined above in a method for preventing HCV infection.

The present invention also relates to a vaccine for immunizing a mammal against HCV infection, comprising at least one peptide as defined above, in a pharmaceutically acceptable carrier.

The present invention relates also to a vaccine as defined above, comprising at least one peptide as defined above, with said peptide being unique for at least one of the subtypes or types as defined in Table 5.

The present invention relates to an antibody raised upon immunization with at least one polypeptide or peptide as defined above, with said antibody being specifically reactive with any of said polypeptides or peptides, and with said antibody being preferably a monoclonal antibody.

The monoclonal antibodies of the invention can be produced by any hybridoma liable to be formed according to classical methods from splenic cells of an animal, particularly from a mouse or rat, immunized against the HCV polypeptides according to the invention as defined above on the one hand, and of cells of a myeloma cell line on the other hand, and to be selected by the ability of the hybridoma to produce the monoclonal antibodies recognizing the polypeptides which has been initially used for the immunization of the animals.

The antibodies involved in the invention can be labelled by an appropriate label of the enzymatic, fluorescent, or radioactive type.

The monoclonal antibodies according to this preferred embodiment of the invention may be humanized versions of mouse monoclonal antibodies made by means of recombinant DNA technology, departing from parts of mouse and/or human genomic DNA sequences coding for H and L chains or from cDNA clones coding for H and L chains.

Alternatively the monoclonal antibodies according to this preferred embodiment of the invention may be human monoclonal antibodies. These antibodies according to the present embodiment of the invention can also be derived from human peripheral blood lymphocytes of patients infected with HCV type 1 subtype 1d, 1e, 1f or 1g, HCV type 2 subtype 2e, 2f, 2g, 2h, 2i, 2k or 2l; HCV type 3, subtype 3g; HCV type 4 subtype 4k, 4l or 4m; and/or HCV type 7 (subtypes 7a, 7c or 7d), 9, 10 or 11, or vaccinated against HCV. Such human monoclonal antibodies are prepared, for instance, by means of human peripheral blood lymphocytes (PBL) repopulation of severe combined immune deficiency (SCID) mice (for recent review, see Duchosal et al. 1992) or by screening Eppstein Barr-virus-transformed lymphocytes of infected or vaccinated individuals for the presence of reactive B-cells by means of the antigens of the present invention.

The invention also relates to the use of the proteins of the invention, muteins thereof, or peptides derived therefrom for the selection of recombinant antibodies by the process of repertoire cloning (Persson et al., 1991).

Antibodies directed to peptides derived from a certain genotype may be used either for the detection of such HCV genotypes, or as therapeutic agents.

The present invention relates also to a method for detecting HCV antigens present in a biological sample, comprising:

(i) contacting said biological sample with an antibody as defined above,

(ii) detecting the immune compleexes formed between said HCV antigens and said antibody.

The present invention relates also to a method for HCV typing, comprising:

(i) contacting said biological sample with an antibody as defined above,

(ii) detecting the immune compleexes formed between said HCV antigens and said antibody.

The present invention relates also to a diagnostic kit for use in detecting the presence of HCV, said kit comprising at least one antibody as defined above, with said antibody being preferably bound to a solid support.

The present invention relates also to a diagnostic kit for HCV typing, said kit comprising at least one antibody as defined above, with said antibody being preferably bound to a solid support.

The present invention relates also to a diagnostic kit as defined above, said kit comprising a range of said antibodies which are attached to specific locations on a solid substrate.

The present invention relates also to a pharmaceutical composition comprising at least one antibody as defined above and a suitable excipient, diluent or carrier.

The present invention relates also to a method of preventing or treating HCV infection, comprising administering the pharmaceutical composition as defined above to a mammal in effective amount.

The present invention relates also to the use of a composition as defined above in a method for preventing or treating HCV infection.

The genotype may also be detected by means of a type-specific antibody as defined above, which may also linked to any polynucleotide sequence that can afterwards be amplified by PCR to detect the immune complex formed (Immuno-PCR, Sano et al., 1992).

Any publications or patent applications referred to herein are incorporated by reference. The following examples illustrate aspects of the invention but are in no way intended to limit the scope thereof.

FIGURE LEGENDS

Figure Legends

Figure 1

Alignment of the nucleotide sequences of the Core/E1 region of some of the isolates of the newly identified types and subtypes of the present invention, with other known prototype isolates of subtypes.

Figure 2

Alignment of the amino acid sequences of the Core/E1 region of some of the isolates of the newly identified types and subtypes of the present invention, with other known prototype isolates of subtypes.

Figure 3

Nucleotide and amino acid sequences obtained from the new HCV isolates of the present invention (SEQ ID NO 1 to 106).

Figure 4

Alignment of the amino acid sequences of the Core/E1 region of some of the isolates of the newly identified types and subtypes of the present invention, with other known prototype isolates of subtypes.

Figure 5

Alignment of the nucleotide sequences of the NS5b region of some of the isolates of the newly identified types and subtypes of the present invention, with other known prototype isolates of subtypes.

Figure 6

Alignment of the amino acid sequences of the NS5b region of some of the isolates of the newly identified types and subtypes of the present invention, with other known prototype isolates of subtypes.

Table 5

Overview of the new subtypes and types of the present invention and the regions sequenced. The subtypes between barckets have been replaced by the non-bracketed subtypes following the classification of Tokita et al. (1994).

Examples

Serum samples.

Serum samples from Cameroonian blood donors (CAM) were screened for HCV antibodies with Innotest HCV Ab III, and confirmed by INNO-LIA HCV III (Innogenetics, Antwerp, Belgium). Serum samples from patients with chronic hepatitis C infection were obtained from various centers in the Benelux countries (BNL), from France (FR), from Pakistan (PAK), from Egypt (EG), and from Vietnam (VN).

Samples from the Benelux, Cameroon, France and Vietnam were selected because of their aberrant reactivities (isolates CAM1078, FR2, FR1, VN4, VN12, VN13, NE98 and others (see Table 5)).

cPCR, LiPA, cloning and sequencing.

RNA isolation, cDNA synthesis, PCR, cloning, and LiPA genotyping using biotinylated 5' UR amplification products were performed as described (Stuyver et al., 1994c). The 5' UR, the Core/E1, and the NS5B PCR products were used for direct sequencing. The sequence of the universal 5' UR primers HCPr95, HCPr96, HCPr98, and HCPr29, were described previously (Stuyver et al. 1993b). The following primers were also described (Stuyver et al. 1994c): HCPr41, a sense

primer for the amplification of the Core region; HCPr52 and HCPr54 for amplification of the Core/E1 region; and HCPr206 and HCPr207 for amplication of a 340-bp NS5B region.

Serum samples BNL1, BNL2, BNL3, BNL4, BNL5, BNL6, BNL7, BNL8, BNL9, BNL10, BNL11, BNL12, CAM1078, FR2, FR16, FR4, FR13, VN13, VN4, VN12, FR1, NE98, and FR19 were analyzed in the Core/E1 region by direct sequencing. Serum samples BNL1, BNL2, FR17, CAM1078, FR2, FR16, BNL3, FR4, BNL5, FR13, FR18, PAK64, BNL8, BNL12, EG81, VN13, VN4, VN12, FR1, NE98, FR14, FR15, and FR19 were also analyzed in the NS5B region by direct sequencing. Partial 5' UR, Core, E1, and NS5B sequences were obtained. The length of the obtained sequences is sufficient to classify the obtained sequences into new types or subtypes, based on the phylogenetic distances to known sequences. The following sequences could be obtained (nucleotide sequences have odd-numbered SEQ ID NO., amino acid sequences have even-numbered SEQ ID NO.): SEQ ID NO 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103 and 105. The amino acid sequences deduced therefrom are given in SEQ ID NO 2, 4, 7, 9, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104 and 106. Table 5 gives an overview of these sequences.

Table 5

Type	Isolate	Nucleotide sequence position		
1d	BNL1	1-310 (SEQ ID NO. 1)	478-925 (SEQ ID NO. 3)	7932-8271 (SEQ ID NO. 53)
1d	BNL2	1-310 (SEQ ID NO. 5)	478-925 (SEQ ID NO. 7)	7932-8271 (SEQ ID NO. 55)
1d	FR17			7932-8271(SEQ ID NO. 57)
1e	CAM1078	1-223 (SEQ ID NO. 9)	(-238)-414 (SEQ ID NO. 59)	7932-8271 (SEQ ID NO. 61)
1f	FR2	1-950 (SEQ ID NO. 11)		7932-8271 (SEQ ID NO. 63)
1g	FR16	(-15)-816 (SEQ ID NO. 65)		7932-8271(SEQ ID NO. 67)
2e	BNL3	1-310 (SEQ ID NO. 13)	478-957 (SEQ ID NO. 15)	7932-8271(SEQ ID NO. 69)
2f	FR4	1-957 (SEQ ID NO 17)		7932-8271(SEQ ID NO. 71)
2g	BNL4		478-925 (SEQ ID NO. 19)	
2h	BNL5	1-310 (SEQ ID NO. 21)	478-925 (SEQ ID NO. 23)	7932-8271(SEQ ID NO. 73)
2i	BNL6		478-833 (SEQ ID NO. 25)	
2k	FR13	(-238)-957 (SEQ ID NO. 75)		7932-8271(SEQ ID NO. 77)
2l	FR18			7932-8271(SEQ ID NO. 79)
3g	PAK64			7932-8271(SEQ ID NO. 81)
4k	BNL7	1-310 (SEQ ID NO. 27)	478-925 (SEQ ID NO. 29)	
4k	BNL8		478-925 (SEQ ID NO. 31)	7932-8271 (SEQ ID NO. 83)
4k	BNL9		478-925 (SEQ ID NO. 33)	
4k	BNL10		478-925 (SEQ ID NO. 35)	
4k	BNL11		478-925 (SEQ ID NO. 37)	
4l	BNL12		478-925 (SEQ ID NO. 39)	7932-8271 (SEQ ID NO. 85)
4m	EG81			7932-8271 (SEQ ID NO. 87)
7a(8b)	VN13	1-413 (SEQ ID NO. 45)		7932-8271 (SEQ ID NO. 89)
7c(8a)	VN4	1-957 (SEQ ID NO. 43)		7932-8271 (SEQ ID NO. 91)
7d(9a)	VN12	1-957 (SEQ ID NO. 47)		7932-8271 (SEQ ID NO. 93)
9a(7a)	FR1	1-957 (SEQ ID NO. 41)		7932-8271 (SEQ ID NO. 95)
10a	NE98	1-310 (SEQ ID NO. 49)	478-925 (SEQ ID NO. 51)	7932-8271 (SEQ ID NO. 97)
11a	FR14			7932-8266 (SEQ ID NO. 99)
11a	FR15			7932-8271(SEQ ID NO. 101)
11a	FR19	(-238)-223 (SEQ ID NO. 103)		7932-8271 (SEQ ID NO. 105)

Table 5-continued

Type	Isolate	Amino acid sequence position		
1d	BNL1	1-103 (SEQ ID NO. 2)	159-308 (SEQ ID NO. 4)	2645-2757 (SEQ ID NO. 54)
1d	BNL2	1-103 (SEQ ID NO. 6)	159-308 (SEQ ID NO. 8)	2645-2757 (SEQ ID NO. 56)
1d	FR17			2645-2757 (SEQ ID NO. 58)
1e	CAM1078	1-74 (SEQ ID NO. 10)	1-138 (SEQ ID NO. 60)	2645-2757 (SEQ ID NO. 62)
1f	FR2	1-316 (SEQ ID NO. 12)		2645-2757 (SEQ ID NO. 64)
1g	FR16	1-158 (SEQ ID NO. 66)		2645-2757 (SEQ ID NO. 68)
2e	BNL3	1-103 (SEQ ID NO. 14)	159-317 (SEQ ID NO. 16)	2645-2757 (SEQ ID NO. 70)
2f	FR4	1-317 (SEQ ID NO. 18)		2645-2757 (SEQ ID NO. 72)
2g	BNL4		159-308 (SEQ ID NO. 20)	
2h	BNL5	1-103 (SEQ ID NO. 22)	159-308 (SEQ ID NO. 24)	2645-2757 (SEQ ID NO. 74)
2i	BNL6		159-277 (SEQ ID NO. 26)	
2k	FR13	1-316 (SEQ ID NO. 76)		2645-2757 (SEQ ID NO. 78)
2l	FR18			2645-2757 (SEQ ID NO. 80)
3g	PAK64			2645-2757 (SEQ ID NO. 82)
4k	BNL7	1-103 (SEQ ID NO. 28)	159-308 (SEQ ID NO. 30)	
4k	BNL8		159-308 (SEQ ID NO. 32)	2645-2757 (SEQ ID NO. 84)
4k	BNL9		159-308 (SEQ ID NO. 34)	
4k	BNL10		159-308 (SEQ ID NO. 36)	
4k	BNL11		159-308 (SEQ ID NO. 38)	
4l	BNL12		159-308 (SEQ ID NO. 40)	2645-2757 (SEQ ID NO. 86)
4m	EG81			2645-2757 (SEQ ID NO. 88)
7a (8b)	VN13	1-137 (SEQ ID NO. 46)		2645-2757 (SEQ ID NO. 90)
7c (8a)	VN4	1-317 (SEQ ID NO. 44)		2645-2757 (SEQ ID NO. 92)
7d (9a)	VN12	1-317 (SEQ ID NO. 48)		2645-2757 (SEQ ID NO. 94)
9a (7a)	FR1	1-317 (SEQ ID NO. 42)		2645-2757 (SEQ ID NO. 96)
10a	NE98	1-103 (SEQ ID NO. 50)	159-308 (SEQ ID NO. 52)	2645-2757 (SEQ ID NO. 98)
11a	FR14			2645-2755 (SEQ ID NO. 100)
11a	FR15			2645-2757 (SEQ ID NO. 102)
11a	FR19	1-74 (SEQ ID NO. 104)		2645-2757 (SEQ ID NO. 106)

Phylogenetic analysis.

Previously published sequences were taken from the EMBL/Genbank database. Alignments were created using the program HCVALIGN (Stuyver et al. 1994c). Sequences were presented in a sequential format to the Phylogeny Inference Package (PHYLIP) version 3.5c (public domain program freely available from the University of Washington, Seattle, USA). Distance matrices were produced by DNADIST using the Kimura 2-parameter setting and further analyzed in NEIGHBOR, using the neighbor-joining setting. The program DRAWTREE was used to create graphic outputs.

Identification of new subtypes

These analyses indicated the clustering of BNL1, BNL2, CAM 1078, FR2, FR16, and FR17 with type 1 isolates, yet neither of these sequences clustered together with any of the known type 1 subtypes 1a, 1b, or 1c. BNL1, BNL2, and FR17 clearly clustered together and could be assigned a new type 1 subtype 1d, while CAM1078 could be classified into another new subtype 1e, FR2 could be classified into another type 1 subtype 1f, and FR16 could be classified into yet another type 1 subtype 1g. Interestingly, all 3 type 1d isolates (BNL1, BNL2, and FR17) and 1g isolate FR16 were obtained from patients of Moroccan ethnic origin who resided in Europe.

Another group of isolates showed homology to other type 2 sequences, but none of the isolates BNL3, FR4, BNL4, BNL5, BNL6, FR13, or FR18 could be classified into one of the known type 2 subtypes 2a, 2b, 2c (Bukh et al., 1993), or 2d (Stuyver et al., 1994c). Based on the phylogenetic distances to other type 2 isolates and to other isolates of the group, each of these isolates could be classified into a new type 2 subtype. BNL3 was assigned subtype 2e, FR4 subtype 2f, BNL4 subtype 2g, BNL5 subtype 2h, and BNL6 could be classified into yet another type 2 subtype 2i. If the previously published isolate HN4 is classified as 2j, FR13 and FR18 may be classified into new type 2 subtypes 2k and 2l. However, the possibility that FR13 and FR18 could belong to subtypes 2g or 2i has not yet been ruled out. Definite classification can be obtained by determining the NS5B sequences of isolates BNL4 and BNL6, belonging to subtypes 2g and 2i, respectively.

Isolate PAK64 showed homology to type 3 sequences, but could not be classified into one of the known type 3 subtypes 3a to f. Based on the phylogenetic distances to other type 3 isolates, PAK64 could be classified into a new type 3

subtype. PAK64 was assigned subtype 3g. However, the possibility that PAK64 belongs to a known type 3 subtype can not be strictly ruled out since only one region of the genome has been sequenced. Definite classification can be obtained by determining the Core/E1 sequences of isolate PAK64 after amplification with primerHcPr52 and HcPr54.

Among the Benelux and Egyptian samples that were analyzed, some sequences clustered with the previously identified type 4 subtypes 4c and 4d. However, BNL7, BNL8, BNL9, BNL10, BNL11, BNL12, and EG81 clustered into new subtypes of type 4. Isolates BNL7, BNL8, BNL9, BNL10, and BNL11 clustered again separately from BNL12 and EG81 into a new subtype 4k. This subtype was the predominant subtype in the Benelux countries. BNL12 and EG81 also segregated into separate subtypes. BNL12 was assigned to another new subtype 4l and EG81 was assigned to yet another new subtype 4m.

Identification of new HCV major types

Isolates FR1, VN4, VN12, VN13, NE98, FR14, FR15, and FR19 did not cluster with any of the known 6 major types of HCV. VN4, VN12, and VN13 were very distantly related to genotype 6, but phylogenetic analysis indicated that these isolates should be assigned new major types. VN13, VN4 and VN12 were related at the subtype level and assigned type 7a, 7c, and 7d, respectively. FR1 was not related to any known isolate and was assigned genotype 9a. NE98 shows a distant relatedness to type 3 sequences, yet phylogenetic analysis suggested classification into a new major type 10a. Depending on international guidelines for assigning type and subtype levels, NE98 may also be classified into an additional type 3 subtype. FR14, FR15, and FR19 show a very distant relatedness to type 2 sequences, yet phylogenetic analysis indicated thes isolates to be classified into a new major type 11, all belonging to the same subtype designated 11a. Depending on international guidelines for assigning type and subtype levels, FR14, FR15, and FR19 may also be classified into an additional type 2 subtype.

REFERENCES

Barany F (1991). Genetic disease detection and DNA amplification using cloned thermostable ligase. Proc Natl Acad Sci USA 88: 189-193.

Bej A, Mahbubani M, Miller R, Di Cesare J, Haff L, Atlas R (1990) Mutiplex PCR amplification and immobilized capture probes for detection of bacterial pathogens and indicators in water. Mol Cell Probes 4:353-365.

Bukh J, Purcell R, Miller R (1992). Sequence analysis of the 5' noncoding region of hepatitis C virus. Proc Natl Acad Sci USA 89:4942-4946.

Bukh J, Purcell R, Miller R (1993). At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. Proc. Natl. Acad. Sci. USA 90,8234-8238.

Cha T, Beal E, Irvine B, Kolberg J, Chien D, Kuo G, Urdea M (1992) At least five related, but distinct, hepatitis C viral genotypes exist. Proc Natl Acad Sci USA 89:7144-7148.

Chan S-W, Simmonds P, McOmish F, Yap P, Mitchell R, Dow B, Follett E (1991) Serological responses to infection with three different types of hepatitis C virus. Lancet 338:1991.

Chan S-W, McOmish F, Holmes E, Dow B, Peutherer J, Follett E, Yap P, Simmonds P (1992) Analysis of a new hepatitis C virus type and its phylogenetic relationship to existing variants. J Gen Virol 73:1131-1141.

Chomczynski P, Sacchi N (1987) Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162:156-159.

Choo Q, Richman K, Han J, Berger K, Lee C, Dong C, Gallegos C, Coit D, Medina-

Selby A, Barr P, Weiner A, Bradley D, Kuo G, Houghton M (1991) Genetic organization and diversity of the hepatitis C virus. Proc Natl Acad Sci USA 88:2451-2455.

Compton J (1991). Nucleic acid sequence-based amplification. Nature, 350: 91-92.

Duchosal A, Eming S, Fisher P (1992) Immunization of hu-PBL-SCID mice and the resue of human monoclonal Fab fragments through combinatorial libraries. Nature 355:258-262.

Duck P (1990). Probe amplifier system based on chimeric cycling oligonucleotides. Biotechniques 9, 142-147.

Guatelli J, Whitfield K, Kwoh D, Barringer K, Richman D, Gengeras T (1990) Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. Proc Natl Acad Sci USA 87: 1874-1878.

Hijikata M, Kato N, Ootsuyama Y, Nakagawa M, Shimotohmo K (1991) Gene mapping of the putative structural region of the hepatitis C virus genome by in vitro processing analysis. Proc Natl Acad Sci USA 88, 5547-5551.

Jacobs K, Rudersdorf R, Neill S, Dougherty J, Brown E, Fritsch E (1988) The thermal stability of oligonucleotide duplexes is sequence independent in tetraalkylammonium salt solutions: application to identifying recombinant DNA clones. Nucl Acids Res 16:4637-4650.

Kato N, Hijikata M, Ootsuyama Y, Nakagawa M, Ohkoshi S, Sugimura T, Shimotohno K (1990) Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. Proc Natl Acad Sci USA 87:9524-9528.

Kwoh D, Davis G, Whitfield K, Chappelle H, Dimichele L, Gingeras T (1989). Transcription-based amplification system and detection of amplified human immunodeficiency virus type 1 with a bead-based sandwich hybridization format. Proc Natl Acad Sci USA, 86: 1173-1177.

Kwok S, Kellogg D, McKinney N, Spasic D, Goda L, Levenson C, Sinisky J, (1990). Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency views type 1 model studies. Nucl. Acids Res., 18: 999.

Landgren U, Kaiser R, Sanders J, Hood L (1988). A ligase-mediated gene detection technique. Science 241:1077-1080.

Lizardi P, Guerra C, Lomeli H, Tussie-Luna I, Kramer F (1988) Exponential amplification of recombinant RNA hybridization probes. Bio/Technology 6:1197-1202.

Lomeli H, Tyagi S, Printchard C, Lisardi P, Kramer F (1989) Quantitative assays based on the use of replicatable hybridization probes. Clin Chem 35: 1826-1831.

Machida A, Ohnuma H, Tsuda F, Munekata E, Tanaka T, Akahane Y, Okamoto H, Mishiro S (1992) Hepatology 16, 886-891.

Maniatis T, Fritsch E, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Mori S, Kato N, Yagyu A, Tanaka T, Ikeda Y, Petchclai B, Chiewsilp P, Kurimura T, Shimotohno K (1992) A new type of hepatitis C virus in patients in Thailand. Biochem Biophys Res Comm 183:334-342.

Okamoto H, Okada S, Sugiyama Y, Kurai K, Iizuka H, Machida A, Miyakawa Y, Mayumi M (1991) Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions. J Gen Virol 72:2697-2704.

Okamoto H, Kurai K, Okada S, Yamamoto K, Lizuka H, Tanaka T, Fukuda S, Tsuda F, Mishiro S (1992) Full-length sequences of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. Virology 188:331-341.

Persson M, Caothien R, Burton D (1991). Generation of diverse high-affinity human monoclonal antibodies by repertoire cloning. Proc Natl Acad Sci USA 89:2432-2436.

Saiki R, Gelfand D, Stoffel S, Scharf S, Higuchi R, Horn G, Mullis K, Erlich H (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487-491.

Saiki R, Walsh P, Levenson C, Erlich H (1989) Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes (1989) Proc Natl Acad Sci USA 86:6230-6234.

Sano T, Smith C, Cantor C (1992) Immuno-PCR: very sensitive antigen detection by means of specific antibody-DNA conjugates. Science 258:120-122.

Simmonds P, McOmsh F, Yap P, Chan S, Lin C, Dusheiko G, Saeed A, Holmes E (1993a), Sequence variability in the 5' non-coding region of hepatitis C virus : identification of a new virus type and restrictions on sequence diversity. J Gen Virology, 74:661-668.

Stuyver L, Rossau R, Wyseur A, Duhamel M, Vanderborght B, Van Heuverswyn H, Maertens G (1993b) Typing of hepatitis C virus (HCV) isolates and characterization of new (sub)types using a Line Probe Assay. J Gen Virology, 74: 1093-1102.

Tokita et al. (1994) Hepatitis C virus vraiants from Vietnam are classifiable into the seventh, eighth, and ninth major genetic groups. Proc. Natl. Acad. Sci, 91: 11022-11026.

Walker G, Little M, Nadeau J, Shank D (1992). Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system. Proc Natl Acad Sci USA 89:392-396.

Wu D, Wallace B (1989). The ligation amplification reaction (LAR) - amplification of specific DNA sequences using sequential rounds of template-dependent ligation. Genomics 4:560-569.

Miller P, Yano J, Yano E, Carroll C, Jayaram K, Ts'o P.(1979) Nonionic nucleic acid analogues. Synthesis and characterization of dideoxyribonucleoside methylphosphonates. Biochemistry 18(23):5134-43.

Nielsen P, Egholm M, Berg R, Buchardt O (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. Science 254(5037):1497-500.

Nielsen P, Egholm M, Berg R, Buchardt O (1993) Sequence specific inhibition of DNA restriction enzyme cleavage by PNA. Nucleic-Acids-Res. 21(2):197-200.

Asseline U, Delarue M, Lancelot G, Toulme F, Thuong N (1984) Nucleic acid-binding molecules with high affinity and base sequence specificity : intercalating agents covalently linked to oligodeoxynucleotides. Proc. Natl. Acad. Sci. USA 81(11):3297-301.

Matsukura M, Shinozuka K, Zon G, Mitsuya H, Reitz M, Cohen J, Broder S (1987) Phosphorothioate analogs of oligodeoxynucleotides : inhibitors of replication and cytopathic effects of human immunodeficiency virus. Proc. Natl. Acad. Sci. USA 84(21):7706-10.

Maertens, G., Ducatteeuw, A., Stuyver, L., Vandeponseele, P., Venneman, A., Wyseur, A., Bosman, F., Heijtink, R. & de Martynoff, G. (1994) Low prevalence of anti-E1 antibodies reactive to recombinant type 1b E1 envelope protein in type 2, 3, and 4 sera, but high prevalence in subtypes 1a and 1b. In: Viral Hepatitis and Liver Disease,, Proceedings of the International Symposium on Viral Hepatitis and Liver Disease (Eds. Nishioka, K., Suzuki, H., Mishiro, S., and Oda, T.), pp 314-316, Springer-Verlag Tokyo.

Simmonds, P., Rose, K.A., Graham, S., Chan, S.-W., McOmish, F., Dow, B.C., Follett, E.A.C., Yap, P.L., & Marsden, H. (1993b) Mapping of serotype-specific, immunodominant epitopes in the NS4 region of hepatitis C virus (HCV): Use of type-specific peptides to serologically discriminate infections with HCV type 1, 2, and 3.

J. Clin. Microbiol. **31**, 1493-1503.

Simmonds, P., Holmes, E.C., Cha, T.-A., Chan, S.-W., McOmish, F., Irvine, B., Beall, E., Yap, P.L., Kolberg, J., & Urdea, M.S. (1993c) *J. Gen. Virol.* **74**, 2391-2399.

Stuyver, L., Van Arnhem, W., Wyseur, A. & Maertens, G. (1994) Cloning and phylogenetic analysis of the Core, E2, and NS3/4 regions of hepatitis C virus type 5a. Biochem. Biophys. Res. Comm. **202**, 1308-1314.

Simmonds, P., Alberti, A., Alter, H., Bonino, F., Bradley, D.W., Bréchot, C., Brouwer, J., Chan, S.-W., Chayama K., Chen, D.-S., Choo, Q.-L., Colombo, M., Cuypers, T., Date, T., Dusheiko, G., Esteban, J.I., Fay, O., Hadziyannis, S., Han, J., Hatzakis, A., Holmes, E.C., Hotta, H., Houghton, M., Irvine, B., Kohara, M., Kolberg, J.A., Kuo, G., Lau, J.Y.N., Lelie, P.N., Maertens, G., McOmish, F., Miyamura, T., Mizokami, M., Nomoto, A., Prince A.M., Reesink, H.W., Rice, C., Roggendorf, M., Schalm, S., Shikata, T., Shimotohno, K., Stuyver, L., Trépo, C., Weiner, A., Yap, P.L. & Urdea, M.S. (1994) A proposed system for the nomenclature of hepatitis C virus genotypes. *Hepatology* **19**, 1321-1324.

Stuyver, L., Van Arnhem, W., Wyseur, A., DeLeys, R. & Maertens, G. (1993a) Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3. Biochem. Biophys. Res. Comm. **192**, 635-641.

Stuyver, L., Rossau, R., Wyseur, A., Duhamel, M., Vanderborght, B., Van Heuverswyn, H. & Maertens, G. (1993b) Typing of hepatitis C virus isolates and characterization of new subtypes using a line probe assay. J. Gen Virol. **74**, 1093-1102.

Stuyver, L., Wyseur, A., Van Arnhem, W., Rossau, R., Delaporte, E., Dazza, M.-C., Van Doorn, L.-J., Kleter, B. & Maertens, G. (1994a) The use of a line probe assay as a tool to detect new types or subtypes of hepatitis C virus. In: Viral Hepatitis and Liver Disease, Proceedings of the International Symposium on Viral Hepatitis and Liver Disease (Eds. Nishioka, K., Suzuki, H., Mishiro, S., and Oda, T.), pp 317-319,

CLAIMS

1. An HCV polynucleic acid, having a nucleotide sequence which is unique to a theretofore unidentified HCV type or subtype which is different from HCV subtypes 1a, 1b, 1c, 2a, 2b, 2c, 2d, 3a, 3b, 3c, 3d, 3e, 3f, 4a, 4b, 4c, 4d, 4e, 4f, 4g, 4h, 4i, 4j, 5a or 6a, with said HCV subtypes being classified as in Table 3 by comparison of a part of the NS5 gene nucleotide sequence spanning positions 7932 to 8271, with said amino acid numbering being shown in Table 1, and with said polynucleic acid containing at least one nucleotide differing from said known HCV nucleotide sequences, or the complement thereof.

2. A polynucleic acid according to claim 1, having a nucleotide sequence which is unique to at least one of HCV subtypes 1d, 1e, 1f, 1g, 2e, 2f, 2g, 2h, 2i, 2k, 2l, 3g, 4k, 4l, 4m, 7a, 7c or 7d, with said HCV subtypes being classified as defined in claim 1.

3. A polynucleic acid according to claim 1, having a nucleotide sequence which is unique to at least one of HCV types 9, 10 or 11, with said HCV types being classified as defined in claim 1.

4. A polynucleic acid according to any of claims 1 to 3 encoding an HCV polyprotein comprising in its amino acid sequence at least one of the following amino acid residues:
I15, C38, V44, A49, Q43, P49, Q55, A58, S60 or D60, E68 or V68, H70, A71 or Q71 or N71, D72, H81, H101, D106, S110, L130, I134, E135, L140, S148, T150 or E150, Q153, F155, D157, G160, E165, I169, F181, L186, T190, T192 or I192 or H192, I193, A195, S196, R197 or N197 or K197, Q199 or D199 or H199 or N199, F200 or T200, A208, I213, M216 or S216, N217 or S217 or G217 or K217, T218, I219, A222, Y223, I230, W231 or L231, S232 or H232 or A232, Q233, E235 or L235, F236 or T236, F237, L240 or M240, A242, N244, N249, I250 or K250 or R250, A252 or C252, A254, I255 or V255, D256 or M256, E257, E260 or K260, R261, V268, S272 or R272, I285, G290 or F290, A291, A293 or L293 or W293, T294 or A294, S295 or H295, K296 or E296, Y297 or M297, I299 or Y299, I300,

Springer-Verlag Tokyo.

Stuyver, L., Van Arnhem, W., Wyseur, A. & Maertens, G. (1994b) Cloning and Phylogenetic analysis of the Core, E2, and NS3/4 regions of the hepatitis C virus type 5a. Biochem. Biophys. Res. Comm. **202**, 1308-1314.

Stuyver, L., Van Arnhem, W., Wyseur, A., Hernandez, F., Delaporte, E., & Maertens, G. (1994c) Classification of hepatitis C viruses based on phylogenetics analysis of the E1 and NS5B regions and identification of 5 new subtypes. Proc. Natl. Acad. Sci. USA **91**.

Stuyver et al. (1995) Hepatitis C virus genotyping by means of 5'-UR/core line probe assays and molecular analysis of untypeable samples. Virus Reasearch (in press).

S301, P316, S2646, A2648, G2649, A2650, V2652, Q2653, H2656 or L2656, D2657, F2659, K2663 or Q2663, A2667 or V1667, D2677, L2681, M2686 or Q2686 or E2686, A2692 or K2692, H2697, I2707, L2708 or Y2708, A2709, A2719 or M2719, F2727, T2728 or D2728, E2729, F2730 or Y2730, I2741, I2745, V2746 or E2746 or L2746 or K2746, A2748, S2749 or P2749, R2750, E2751, D2752 or N2752 or S2752 or T2752 or V2752 or I2752 or Q2752, S2753 or D2753 or G2753, D2754, A2755, L2756 or Q2756, R2757,

with said notation being composed of a letter representing the amino acid residue by its one-letter code, and a number representing the amino acid numbering as shown in Table 1,

or a part of said polynucleic acid which is unique to at least one of the HCV subtypes or types as defined in claims 2 to 3, and which contains at least one nucleotide differing from known HCV nucleotide sequences, or the complement thereof.

5. A polynucleic acid according to any of claims 1 to 4, with said polynucleic acid encoding a HCV polyprotein comprising in its amino acid sequence at least one amino acid sequence chosen from the following list:

ARQSDGRSWAQ or ARRSEGRSWAQ as for subtype 1d	(SEQ ID NO 107 and 108)
ERRPEGRSWAQ as for subtype 1e	(SEQ ID NO 109)
ARRPEGRSWAQ as for subtype 1f	(SEQ ID NO 110)
DRRTTGKSWGR as for subtype 2k	(SEQ ID NO 111)
DRRATGRSWGR as for subtype 2e	(SEQ ID NO 112)
DRRATGKSWGR as for subtype 2f	(SEQ ID NO 113)
VRQPTGRSWGQ as for type 9	(SEQ ID NO 114)
VRHQTGRTWAQ as for subtype 7a and 7c	(SEQ ID NO 115)
VRQNQGRTWAQ as for subtype 7d	(SEQ ID NO 116)
ARRTEGRSWAQ as for type 10	(SEQ ID NO 117)
VRRTTGRXXXX or VRRTTGRTWAQ as for type 11	(SEQ ID NO 118 and 119)
HEVRNASGVYHV or HEVRNASGVYHL as for subtype 1d	(SEQ ID NO 120 and 121)
YEVHSTTDGYHV as for subtype 1f	(SEQ ID NO 122)
VEVKNTSQAYMA as for subtype 2e	(SEQ ID NO 123)
IQVKNNSHFYMA as for subtype 2f	(SEQ ID NO 124)

VQVKNTSTMYMA as for subtype 2g (SEQ ID NO 125)

VQVKNTSHSYMV as for subtype 2h (SEQ ID NO 126)

VQVANRSGSYMV as for subtype 2i (SEQ ID NO 127)

VEIKNTXNTYVL or VEIKNTSNTYVL as for subtype 2k (SEQ ID NO 128 and 129)

INYRNVSGIYYV or INYRNTSGIYHV or INYHNTSGIYHI or TNYRNVSGIYHV as for subtype 4k (SEQ ID NO 130, 131, 132 or 133)

QHYRNVSGIYHV as for subtype 4l (SEQ ID NO 134)

IQVKNASGIYHL as for type 9 (SEQ ID NO 135)

AHYTNKSGLYHL as for subtype 7c (SEQ ID NO 136)

LNYANKSGLYHL as for subtype 7d (SEQ ID NO 137)

LEYRNASGLYMV as for type 10 (SEQ ID NO 138)

IYEMDGMIMHY or IYEMSGMILHA as for subtype 1d (SEQ ID NO 139 and 140)

VYEAKDIILHT as for subtype 1f (SEQ ID NO 141)

VWQLXDAVLHV as for subtype 2e (SEQ ID NO 142)

VWQLRDAVLHV as for subtype 2f (SEQ ID NO 143)

IWQMQGAVLHV as for subtype 2g (SEQ ID NO 144)

VWQLKDAVLHV as for subtype 2h (SEQ ID NO 145)

VWQLEEAVLHV as for subtype 2i (SEQ ID NO 146)

TWQLXXAVLHV as for subtype 2k (SEQ ID NO 147)

VYEADHHILHL or VYEADHHILAL or VFEADHHILHL as for subtupe 4k (SEQ ID NO 148, 149 and 150)

VYESDHHILHL as for subtype 4l (SEQ ID NO 151)

VFEAETMILHL as for type 9 (SEQ ID NO 152)

VYEAETLILHL as for subtype 7c (SEQ ID NO 153)

VYEANGMILHL as for subtype 7d (SEQ ID NO 154)

VYEAGDIILHL as for type 10 (SEQ ID NO 155)

VREDNHLRCWMAL or VRENNSSRCWMAL as for subtype 1d (SEQ ID NO 156 and 157)

IREGNISRCWVPL as for subtype 1f (SEQ ID NO 158)

ENSSGRFHCWIPI as for subtype 2e (SEQ ID NO 159)

ERSGNRTFCWTAV as for subtype 2f (SEQ ID NO 160)

ELQGNKSRCWIPV as for subtype 2g (SEQ ID NO 162)

ERHQNQSRCWIPV as for subtype 2h (SEQ ID NO 163)

EWKDNTSRCWIPV as for subtype 2i (SEQ ID NO 164)

EREGNSSRCWIPV as for subtype 2k (SEQ ID NO 165)

VREGNQSRCWVAL or VRTGNQSRCWVAL or VRVGNQSSCWVAL or VRVGNQSRCWVAL or VKEGNHSRCWVAL as for subtype 4k (SEQ ID NO 166, 167, 168 or 169)

VKTGNTSRCWVAL as for subtype 4l (SEQ ID NO 170)

IKAGNESRCWLPV as for type 9 (SEQ ID NO 171)

VKEGNQSRCWVQA as for subtype 7c (SEQ ID NO 172)

VKXXNLTKCWLSA as for subtype 7d (SEQ ID NO 173)

VRSGNTSRCWIPV as for type 10 (SEQ ID NO 174)

VKNASVPTAA or VKDANVPTAA as for subtype 1d (SEQ ID NO 175 and 176)

ARIANAPIDE as for subtype 1f (SEQ ID NO 177)

VSKPGALTKG as for subtype 2e (SEQ ID NO 178)

VSRPGALTRG as for subtype 2f (SEQ ID NO 179)

VNQPGALTRG as for subtype 2g (SEQ ID NO 180)

VSQPGALTRG as for subtype 2h (SEQ ID NO 181)

VSQPGALTKG as for subtype 2i (SEQ ID NO 182)

VSRPGALTEG as for subtype 2k (SEQ ID NO 183)

APYIGAPLES or APYTAAPLES as for subtype 4k (SEQ ID NO 184 and 185)

APILSAPLMS as for subtype 4l (SEQ ID NO 186)

VPNSSVPIHG as for type 9 (SEQ ID NO 187)

VPNASTPVTG as for subtype 7c (SEQ ID NO 188)

VQNASVSIRG as for subtype 7d (SEQ ID NO 189)

VKSPCAATAS as for type 10 (SEQ ID NO 190)

SPRMHHTTQE or SPRLYHTTQE as for subtype 1d (SEQ ID NO 191 and 192)

TSRRHWTVQD as for subtype 1f (SEQ ID NO 193)

APKRHYFVQE as for subtype 2e (SEQ ID NO 194)

SPQYHTFVQE as for subtype 2f (SEQ ID NO 195)

SPQHHNFSQD as for subtype 2g (SEQ ID NO 196)

SPQHHIFVQD as for subtype 2h (SEQ ID NO 197)

SPEHHHFVQD as for subtype 2k (SEQ ID NO 198)

RPRRHWTTQD or RPRRHWTAQD or QPRRHWTTQD or RPRRHWTTQE as for

subtype 4k (SEQ ID NO 199, 200, 201 or 202)
QPRRHWTVQD as for subtype 4l (SEQ ID NO 203)
RPKYHQVTQD as for type 9 (SEQ ID NO 204)
RPRMHQVVQE as for subtype 7c (SEQ ID NO 205)
RPRMYEIAQD as for subtype 7d (SEQ ID NO 206)
RHRQHWTVQD as for type 10 (SEQ ID NO 207)

or a part of said polynucleic acid which is unique to at least one of the HCV subtypes or types as defined in claims 2 to 3, and which contains at least one nucleotide differing from known HCV nucleotide sequences, or the complement thereof.

6. A polynucleic acid according to any of claims 1 to 5 having a sequence selected from any of SEQ ID NO 1 to 105, or a part of said polynucleic acid which is unique to at least one of the HCV subtypes or types as defined in claims 2 to 3, and which contains at least one nucleotide differing from known HCV nucleotide sequences, or the complement thereof.

7. A polynucleic acid according to any of claims 1 to 6, which codes for the 5' UR, the Core/E1, the NS4 or the NS5B region or a part thereof.

8. A polynucleic acid according to any of claims 1 to 7 which is a cDNA sequence.

9. An oligonucleotide primer comprising part of a polynucleic acid according to any of claims 1 to 8, with said primer being able to act as primer for specifically amplifying the nucleic acid of a certain isolate belonging to the genotype from which the primer is derived.

10. An oligonucleotide probe comprising part of a polynucleic acid according to any of claims 1 to 8, with said probe being able to act as a hybridization probe for specific detection and/or classification into types and/or subtypes of a HCV nucleic acid containing said nucleotide sequence, with said probe being possibly labelled or attached to a solid substrate.

11. A diagnostic kit for use in determing the genotype of HCV, said kit comprising a

primer according to claim 9.

12. A diagnostic kit for use in determining the genotype of HCV, said kit comprising a probe according to claim 10.

13. A diagnostic kit according to claim 12, wherein said probe(s) is(are) attached to a solid substrate.

14. A diagnostic kit according to claim 13, wherein a range of said probes are attached to specific locations on a solid substrate.

15. A diagnostic kit according to claim 14, wherein said solid support is a membrane strip and said probes are coupled to the membrane in the form of parallel lines.

16. A method for the detection of HCV nucleic acids present in a biological sample, comprising:

(i) possibly extracting sample nucleic acid,

(ii) amplifying the nucleic acid with at least one primer according to claim 9,

(iii) detecting the amplified nucleic acids.

17. A method for the detection of HCV nucleic acids present in a biological sample, comprising:

(i) possibly extracting sample nucleic acid,

(ii) possibly amplifying the nucleic acid with at least one primer according to claim 9, or with a universal HCV primer,

(iii) hybridizing the nucleic acids of the biological sample, possibly under denatured conditions, at appropriate conditions with one or more probes according to claim 10, with said probes being possibly attached to a solid substrate,

(iv) possibly washing at appropriate conditions,

(v) detecting the hybrids formed.

18. A method for detecting the presence of one or more HCV genotypes present in

a biological sample, comprising:

(i) possibly extracting sample nucleic acid,

(ii) specifically amplifying the nucleic acid with at least one primer according to claim 9,

(iii) detecting said amplified nucleic acids,

(iv) inferring the presence of one or more genotypes of HCV present from the observed pattern of amplified fragments.

19. A method for detecting the presence of one or more HCV genotypes present in a biological sample, comprising:

(i) possibly extracting sample nucleic acid,

(ii) possibly amplifying the nucleic acid with at least one primer according to claim 9 or with a universal HCV primer,

(iii) hybridizing the nucleic acids of the biological sample, possibly under denatured conditions, at appropriate conditions with one or more probes according to claim 10, with said probes being possibly attached to a solid substrate,

(iv) possibly washing at appropriate conditions,

(v) detecting the hybrids formed,

(vi) inferring the presence of one or more HCV genotypes present from the observed hybridization pattern.

20. A method according to claim 19, wherein said probes are further characterized as defined in any of claims 13 to 15.

21. A method according to claims 16 to 18, wherein said nucleic acids are labelled during or after amplification.

22. A polypeptide having an amino acid sequence encoded by a polynucleic acid according to any of claims 1 to 8, or a part thereof which is unique to at least one of the HCV subtypes or types as defined in claims 2 or 3, and which contains at least one amino acid differing from any of the known HCV types or subtypes amino acid sequences, or an analog thereof being substantially homologous and biologically

equivalent.

23. A polypeptide according to claim 22 comprising in its amino acid sequence at least one of the following amino acid residues:
I15, C38, V44, A49, Q43, P49, Q55, A58, S60 or D60, E68 or V68, H70, A71 or Q71 or N71, D72, H81, H101, D106, S110, L130, I134, E135, L140, S148, T150 or E150, Q153, F155, D157, G160, E165, I169, F181, L186, T190, T192 or I192 or H192, I193, A195, S196, R197 or N197 or K197, Q199 or D199 or H199 or N199, F200 or T200, A208, I213, M216 or S216, N217 or S217 or G217 or K217, T218, I219, A222, Y223, I230, W231 or L231, S232 or H232 or A232, Q233, E235 or L235, F236 or T236, F237, L240 or M240, A242, N244, N249, I250 or K250 or R250, A252 or C252, A254, I255 or V255, D256 or M256, E257, E260 or K260, R261, V268, S272 or R272, I285, G290 or F290, A291, A293 or L293 or W293, T294 or A294, S295 or H295, K296 or E296, Y297 or M297, I299 or Y299, I300, S301, P316, S2646, A2648, G2649, A2650, V2652, Q2653, H2656 or L2656, D2657, F2659, K2663 or Q2663, A2667 or V2667, D2677, L2681, M2686 or Q2686 or E2686, A2692 or K2692, H2697, I2707, L2708 or Y2708, A2709, A2719 or M2719, F2727, T2728 or D2728, E2729, F2730 or Y2730, I2741, I2745, V2746 or E2746 or L2746 or K2746, A2748, S2749 or P2749, R2750, E2751, D2752 or N2752 or S2752 or T2752 or V2752 or I2752 or Q2752, S2753 or D2753 or G2753, D2754, A2755, L2756 or Q2756, or R2757,
with said notation being composed of a letter representing the amino acid residue by its one-letter code, and a number representing the amino acid numbering as shown in Table 1,
or a part of said polypeptide which is unique to at least one of the HCV subtypes or types as defined in claims 2 to 3, and which contains at least one amino acid differing from known HCV types or subtypes amino acid sequences, or an analog thereof being substantially homologous and biologically equivalent to said polypeptide.

24. A polypeptide according to claim 22 comprising in its amino acid sequence at least one of the sequences represented by SEQ ID NO 107 to 207 as listed in claim 5, or part of said polypeptide which is unique to at least one of the HCV subtypes or types as defined in claims 2 to 3, and which contains at least one amino acid differing from

known HCV types or subtypes amino acid sequences, or an analog thereof being substantially homologous and biologically equivalent to said polypeptide.

25. A polypeptide having an amino acid sequence as represented in any of SEQ ID NO 1 to 106, or a part thereof which is unique to at least one of the HCV subtypes or types as defined in claims 2 to 3, and which contains at least one amino acid differing from known HCV types or subtypes amino acid sequences, or an analog thereof being substantially homologous and biologically equivalent to said polypeptide.

26. A recombinant polypeptide encoded by a polynucleic acid according to any of claims 1 to 8, or a part thereof which is unique to at least one of the HCV subtypes or types as defined in claims 2 or 3, and which contains at least one amino acid differing from known HCV types or subtypes amino acid sequences, or an analog thereof being substantially homologous and biologically equivalent to said polypeptide.

27. A method for production of a recombinant polypeptide of claim 26, comprising:

- transformation of an appropriate cellular host with a recombinant vector, in which a polynucleic acid or a part thereof according to any of claims 1 to 8 has been inserted under the control of the appropriate regulatory elements,
- culturing said transformed cellular host under conditions enabling the expression of said insert, and,
- harvesting said polypeptide.

28. A recombinant expression vector comprising a polynucleic acid or a part thereof according to any of claims 1 to 8 operably linked to prokaryotic, eukaryotic or viral transcription and translation control elements.

29. A host cell transformed with a recombinant vector according to claim 28.

30. A method for detecting antibodies to HCV present in a biological sample, comprising:

(i) contacting the biological sample to be analysed for the presence of HCV with a polypeptide according to any of claims 22 to 26,

(ii) detecting the immunological complex formed between said antibodies and said polypeptide.

31. A method for HCV typing, comprising:

(i) contacting the biological sample to be analysed for the presence of HCV with a polypeptide according to any of claims 22 to 26,

(ii) detecting the immunological complex formed between said antibodies and said polypeptide.

32. A diagnostic kit for use in detecting the presence of HCV, said kit comprising at least one polypeptide according to any of claims 22 to 26, with said polypeptide being possibly bound to a solid support.

33. A diagnostic kit for HCV typing, said kit comprising at least one polypeptide according to any of claims 22 to 26, with said polypeptide being possibly bound to a solid support.

34. A diagnostic kit according to claims 32 to 33, said kit comprising a range of polypeptides which are attached to specific locations on a solid substrate.

35. A diagnostic kit according to claims 32 to 34, wherein said solid support is a membrane strip and said polypeptides are coupled to the membrane in the form of parallel lines.

36. A pharmaceutical composition comprising at least one polypeptide according to any of claims 22 to 26 and a suitable excipient, diluent or carrier.

37. A method of preventing HCV infection, comprising administering the pharmaceutical compositon of claim 36 to a mammal in effective amount to stimulate the production of protective antibody or protective T-cell response.

38. Use of a composition according to claim 36 in a method for preventing HCV infection as defined in claim 37.

39. A vaccine for immunizing a mammal against HCV infection, comprising at least one polypeptide according to claims 22 to 26, in a pharmaceutically acceptable carrier.

40. A vaccine according to claim 39, comprising at least one polypeptide according to claims 22 to 26, with said polypeptide being unique for at least one of the HCV subtypes as defined in claims 2 or 3.

41. A peptide corresponding to an amino acid sequence encoded by at least one of the HCV polynucleic acids according to any of claims 1 to 8, with said peptide comprising an epitope being unique to at least one of the HCV subtypes or types as defined in claims 2 or 3, and with said peptide containing at least one amino acid differing from any of the known HCV types or subtypes amino acid sequences, or an analog thereof being substantially homologous and biologically equivalent.

42. A method for detecting antibodies to HCV present in a biological sample, comprising:

(i) contacting the biological sample to be analysed for the presence of HCV with a peptide according to claim 41,

(ii) detecting the immune complex formed between said antibodies and said peptide.

43. A method for HCV typing, comprising:

(i) contacting the biological sample to be analysed for the presence of HCV with a peptide according to claim 41,

(ii) detecting the immune complex formed between said antibodies and said peptide.

44. A. diagnostic kit for use in detecting the presence of HCV, said kit comprising at least one peptide according to claim 41, with said peptide being possibly bound to a solid support.

45. A diagnostic kit for HCV typing, said kit comprising at least one peptide according to any of claim 41, with said peptide being possibly bound to a solid support.

46. A diagnostic kit according to claims 44 or 45, wherein said peptides are selected from the following list:

- at least one NS4 peptide,
- at least one NS4 peptide and at least one Core peptide,
- at least one NS4 peptide and at least one Core peptide and at least one E1 peptide, or,
- at least one NS4 peptide and at least one E1 peptide.

47. A Diagnostic kit according to claims 44 to 46, said kit comprising a range of peptides which are attached to specific locations on a solid substrate.

48. A diagnostic kit according to claims 44 to 47, wherein said solid support is a membrane strip and said peptides are coupled to the membrane in the form of parallel lines.

49. A pharmaceutical composition comprising at least one peptide according to claim 41 and suitable excipient, diluent or carrier.

50. A method of preventing HCV infection, comprising administering the pharmaceutical composition of claim 49 to a mammal in effective amount to stimulate the production of protective antibody or protective T-cell response.

51. Use of a composition according to claim 49 in a method for preventing HCV infection as defined in claim 50.

52. A vaccine for immunizing a mammal against HCV infection, comprising at least one peptide according to claim 41, in a pharmaceutically acceptable carrier.

53. A vaccine according to claim 52, comprising at least one peptide according to claim 41, with said peptide being unique for at least one of the subtypes or types as defined in claims 2 or 3.

54. An antibody raised upon immunization with at least one polypeptide or peptide

according to any of claims 22 to 26 or 41, with said antibody being specifically reactive with any of said polypeptides or peptides, and with said antibody being preferably a monoclonal antibody.

55. A method for detecting HCV antigens present in a biological sample, comprising:

(i) contacting said biological sample with an antibody according to claim 54,

(ii) detecting the immune complexes formed between said HCV antigens and said antibody.

56. A method for HCV typing, comprising:

(i) contacting said biological sample with an antibody according to claim 54,

(ii) detecting the immune complexes formed between said HCV antigens and said antibody.

57. A diagnostic kit for use in detecting the presence of HCV, said kit comprising at least one antibody according to claim 54, with said antibody being possibly bound to a solid support.

58. A diagnostic kit for HCV typing, said kit comprising at least one antibody according to claim 54, with said antibody being possibly bound to a solid support.

59. A diagnostic kit according to claims 57 to 58, said kit comprising a range of antibodies which are attached to specific locations on a solid substrate.

60. A pharmaceutical composition comprising at least one antibody according to claim 54 and a suitable excipient, diluent or carrier.

61. A method of preventing or treating HCV infection, comprising administering the pharmaceutical composition of claim 62 to a mammal in effective amount.

62. Use of a composition according to claim 60 in a method for preventing or treating HCV infection as defined in claim 61.